

7 May 2010 | \$10

Science



The Neandertal Genome

 AAAS



Brooks Hanson is Deputy Editor for physical sciences at *Science*.

Stepping Back; Moving Forward

THE CONTROVERSIAL E-MAILS RELATED TO CLIMATE CHANGE, PLUS REPORTED ERRORS IN THE Intergovernmental Panel on Climate Change (IPCC) reports, have spurred a dangerous deterioration in the rational relation between science and society. One U.S. senator has called 17 prominent climate scientists criminals, and pundits have suggested that climate scientists should commit suicide. Fourteen U.S. states have filed lawsuits opposing the federal regulation of greenhouse gas emissions, some asserting that “climate change science is a conspiracy.” South Dakota even resolved that there are other “astrological” forcings on climate. Scientists have been barraged by hateful e-mails. The debate has become polarized, and the distrust of scientists and their findings extends well beyond climate science. What can be done to repair society’s trust in science? A broader perspective is needed on all sides.

The main societal challenges—global energy supply, growing the food supply, and improving public health, among others—depend intimately on science, and for this reason society requires a vigorous scientific enterprise. Our expanding global economy is taxing resources and the environment in ways that cannot be sustained. Science provides a deep understanding of these impacts and, as a result, the ability to predict consequences and assess risks.

Addressing anthropogenic climate change exemplifies the challenges inherent in providing critical scientific advice to society (see the Policy Forum on p. 695 and Letter on p. 689). Climate is as global as today’s economy; we know from archaeological and historical records that an unstable climate has disrupted societies. For these reasons, scientists and governments are jointly committed to understanding the impacts of climate change. Thousands of scientists have volunteered for the IPCC or other assessments. Governments have a vested interest in the success of these assessments, and the stakes are high.

We thus must move beyond polarizing arguments in ways that strengthen this joint commitment. The scientific community must recognize that the recent attacks stem in part from its culture and scientists’ behavior. In turn, it is time to focus on the main problem: The IPCC reports have underestimated the pace of climate change while overestimating societies’ abilities to curb greenhouse gas emissions.

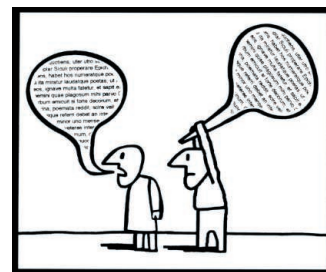
Scientists must meet other responsibilities. The ability to collect, model, and analyze huge data sets is one of the great recent advances in science and has made possible our understanding of global impacts. But developing the infrastructure and practices required for handling data, and a commitment to collect it systematically, have lagged. Scientists have struggled to address standardizing, storing, and sharing data, and privacy concerns. Funding must be directed not only toward basic science but toward facilitating better decisions made with the data and analyses that are produced. As a start, research grants should specify a data curation plan, and there should be a greater focus on long-term monitoring of the environment.

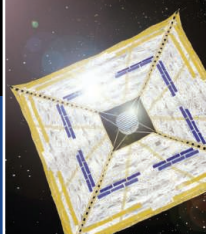
Because society’s major problems are complex, generating useful scientific advice requires synthesizing knowledge from diverse disciplines. As the need for synthesis grows, the avenues of communication are changing rapidly. Unfortunately, many news organizations have eviscerated their science staffs. As a result, stories derived from press releases on specific results are crowding out the thoughtful syntheses that are needed.

If the scientific community does not aggressively address these issues, including communicating its process of discovery and recognizing its modern data responsibilities, and if society does not constructively engage science, then the scientific enterprise and the whole of society are in danger of losing their crucial rational relationship. Carl Sagan’s warnings are especially apt today: “We live in a society exquisitely dependent on science and technology, in which hardly anyone knows anything about science and technology.” “This is a prescription for disaster. We might get away with it for a while, but sooner or later this combustible mixture of ignorance and power is going to blow up in our faces.”

— Brooks Hanson

10.1126/science.1190790





GULF OIL SPILL

Will *Deepwater Horizon* Set a New Standard for Catastrophe?

The fiery destruction of an oil drilling platform in the Gulf of Mexico on 20 April may have triggered one of the worst environmental disasters in U.S. history. The impact of the crisis, which began with the deaths of 11 workers and then simmered for several days before an expanding oil slick grabbed worldwide attention, promises to test the federal government's ability to protect habitat, wildlife, and the economic well-being of a four-state region on a scale never before imagined.

Secretary Ken Salazar of the Department of the Interior declared on 2 May that the government's role is to "keep its boot on the neck of BP," the British-based oil giant that had contracted to use the drilling rig, *Deepwater Horizon*. And BP CEO Tony Hayward said his company "will absolutely be paying for the cleanup operation. There is no doubt about that. It's our responsibility—we we accept it fully."

As *Science* went to press, severe weather was slowing the spread of the oil to the shore. BP officials were hoping to deploy an oil-topping containment vessel in the coming days. Biologists have their fingers crossed, fretting over a blow at the worst possible time to an already fragile ecosystem. "I don't want to be alarmist, but I'm alarmed," says ornithologist James Remsen of Louisiana State University (LSU), Baton Rouge. As a precautionary step to ensure the safety of the food supply, the National Oceanic and Atmospheric Administration (NOAA) has imposed a 10-day ban on commercial fishing in the region.

What went wrong

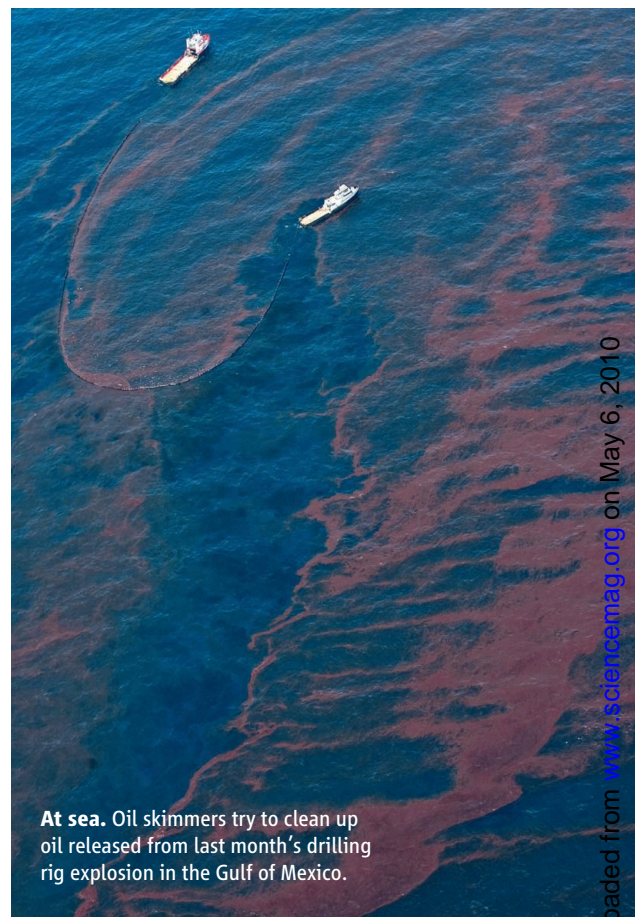
The reasons the drilling vessel failed to contain the 100-million-pascal pressures in the well it had drilled 5500 meters into the sea floor, at a water depth of 1500 meters, are still unknown. But it probably wasn't because its crew pushed too hard on the envelope of modern technology. Just 8 months earlier, in the Gulf of Mexico, *Deepwater Horizon* safely drilled the world's deepest well through 9424 meters of rock in 1259 meters of water.

The latest drilling, 65 kilometers off the Louisiana coast, targeted a shallower, well-known oil-bearing stratum. That made for "rather routine" drilling, says petroleum engineer Kenneth Gray of the University of Texas, Austin. By 20 April, workers had finished drilling the hole using the weight of heavy drilling "mud" filling the hole to contain the pressure. A steel pipe or casing was inserted partway down the hole to reinforce it against collapse before cement was pumped between the casing and the wall of the hole.

This "cement job" is often the culprit in oil spills. According to a 2007 report from the U.S. Minerals Management Service, the agency that supervises offshore drilling, cement-related problems caused 18 of 39 blowouts in the Gulf of Mexico over a 14-year period. Something about the design or execution of the cement job could have prevented the cement from setting properly. That flaw could have allowed oil and gas to work its way upward until containment failed entirely.

Estimates of the flow started at zero but quickly rose to 1000 barrels and then 5000 barrels per day. Before drilling began, BP had projected a worst-case scenario of 162,000 barrels escaping per day. Why that hasn't happened is anyone's guess. The answer could lie in constriction in the well, in the blowout preventer sitting at the top of the well, or in the pipe that once connected the well and the drilling vessel now lying crumpled on the sea floor.

The duration of the spill is another open question. The blowout preventer, which failed to cut off the initial surge, has so far refused to respond to remotely operated submersible vehicles. In a week or two, workers will attempt to capture much of what is still leaking out by lowering a 70-ton concrete-and-metal box over the wellhead leak site. The captured oil would then be pumped out the top to wait-



At sea. Oil skimmers try to clean up oil released from last month's drilling rig explosion in the Gulf of Mexico.

ing barges. No one has ever attempted such a maneuver at 1500-meter depths.

The ultimate solution would involve drilling back into the pressurized rock that's causing all the trouble. The approach requires drilling a new hole from a floating platform, intersecting the original 20-centimeter-wide well kilometers beneath the sea floor, and then pumping in drilling mud to kill the flow. It's an ambitious project that would take at least 3 months.

Dealing with the oil

The typical southern Louisiana crude rising from the well emulsifies into a foamy "mousse" by the time it reaches the surface, and the more volatile, more toxic components quickly evaporate. In the current spill, "Mother Nature is helping us," says marine chemist Edward Overton, professor emeritus at LSU Baton Rouge and head of a chemical hazards assessment team under contract to NOAA. "The possibilities are



Neandertals
and us

680



The mysterious towers
of the Himalayas

685

horrible, but the probabilities are less so.”

The one sample Overton and colleagues have been able to analyze “looked like roof tar,” he says, and such goo is expected to eventually form tarballs and tar mats. “We won’t see a ‘black tide’ in this spill as we did after *Exxon Valdez*,” he predicts. The tarballs “are going to be very sticky but not very toxic. ... We’re looking at an all-summer event.”

Rough seas have stymied the government’s ability to skim and collect the oil on the surface, and hindered some floating protective booms from being put into place properly. Several small fires were used to reduce the total amount of oil on the surface. Once the oil slick began moving toward the coast, BP started the aerial spraying of dispersants—detergent-like chemicals designed to break up the oil. Then engineers added dispersants more directly to where the oil came out of the well. The aim is to reduce the risk to seabirds and mammals, and to protect sensitive habitats from being soaked in oil. But the technology has a side effect: By spreading the oil throughout the water column, fish such as spawning bluefin tuna that would otherwise not be exposed are put at risk of toxic effects.

One possible long-term weapon is to use microbes already present in the water. Another technique, adding fertilizer to coastal waters to enhance the ability of natural microbes to metabolize the oil, has shown promise in small field tests, said Ken Lee of the Bedford Institute of Oceanography in Nova Scotia, Canada. A third approach would involve burning oil-covered wetlands in a way that preserves the roots of the plants.

A pressing economic issue for the region is ensuring the safety of shrimp, mussels, and clams that grow in gulf waters. Generally, local oil spills lead the government to shut down nearby fishing or seafood harvesting areas, and they are reopened after the catch is tested and found to be safe. But fishers lack easy screening tools to help them estimate contaminant levels in their catch, and there are no industry-wide techniques to do the tests accurately and quickly. “We don’t know what [constitutes] a toxic dose,” says former Food and Drug Administration regulator David Acheson. The challenge, he adds, is measuring long-term rather than acute exposure to oil-tainted seafood.

The potential impact

The biggest biological impacts of the spill are expected to be on islands that host large colonies of breeding birds and in the rich coastal wetlands, which nourish young fish, shrimp, and shellfish. “The area of the blowout is one of the most productive areas of the Gulf of Mexico,” says Gilbert Rowe, a marine biologist at Texas A&M University, Galveston.

Scientists are especially concerned about the Breton National Wildlife Refuge. The many breeding seabirds on the barrier islands include some 2000 brown pelicans, which only last year had recovered enough to be removed from the federal list of endangered species. If the oil can’t be kept off the beaches, an entire generation of pelicans and other species could be wiped out, says LSU’s Remsen.

Oil harms seabirds by allowing water to soak through their feathers, chilling the birds and making it harder to gather food. Dolphins are also liable to suffer organ damage from breathing toxic vapors or ingesting too much oil. Sea turtles are at particularly high risk because they feed at the oil-slicked surface.

The wetlands themselves are fragile, too. Having been damaged for decades by dredging, drilling, and other human impacts, Louisiana’s wetlands are disappearing at a rate of 4400 hectares per year (*Science*, 25 November 2005, p. 1264). Oil can harm the vital grasses, which keep the sediment from eroding, by preventing them from photosynthesizing or exchanging gases. Plants cope by sending out new leaves, up to a point. So researchers are holding their breath to see the impact, especially for the most sensitive wetlands farther inland. If the leak isn’t plugged, they could be hard-hit by tropical storms blowing oil inland this summer. “We may be talking catastrophic [effects],” says plant ecologist David White of Loyola University in New Orleans.

One obvious need is for additional research to track the impact of the oil on existing study sites. “We’re poised to do lots of follow-up studies,” says ecologist Denise Reed of the University of New Orleans. “It’s pretty clear that quite a lot of oil is going to get into the wetlands.”

—RICHARD KERR, ELI KINTISCH,
AND ERIK STOKSTAD

Three Historic Blowouts

The decade from 1969 to 1979 witnessed three massive spills from offshore oil wells around the world. Here is how they compare in size and impact.

IXTOC 1

The biggest well-related spill was triggered on 3 June 1979, when a lack of drilling mud allowed oil and gas to shoot up through the 3.6-km-deep IXTOC 1 exploratory well, about 80 km offshore in the southern Gulf of Mexico. The initial daily outflow of 30,000 barrels of oil was eventually reduced to 10,000 barrels. The well was finally capped more than 9 months later. Mexico’s state-owned oil company, PEMEX, treated the approximately 3.5-million-barrel spill with dispersants. U.S. officials had a 2-month head start to reduce impacts to the Texas coastline.



Mexico
1979

Ekofisk

The first major spill in the North Sea resulted in the release of 202,000 barrels of oil about 250 km off the coast of Norway. The 22 April 1977 blowout caused oil to gush from an open pipe 20 m above the sea surface. The well was capped after a week. Between 30% and 40% of the spill evaporated almost immediately. Rough waters broke up the slick before it reached shore.



North Sea
1977

Santa Barbara

A blown well 1 km below the sea floor and 9 km off the coast of Santa Barbara, California, spewed out a total of 100,000 barrels of oil. The initial eruption occurred on 28 January 1969, and the well was capped by mud and cement on 7 February, but the pressure forced oil through sea floor fissures until December. The oil contaminated 65 km of coastline. At least 3700 birds are known to have died, and commercial fishing in the area was closed until April.



Santa Barbara
1969

—LAUREN SCHENKMAN

BIOMEDICAL RESEARCH

Peering Over a Cliff at the Poststimulus World

A year ago, the mood in the U.S. biomedical community was euphoric as researchers scrambled for a piece of the National Institutes of Health's (NIH's) stimulus windfall. More than \$8 billion went to extramural research (and \$1.8 billion for construction and equipment); thousands of scientists are now toiling away on the studies it funded. But reality is setting in. The 2-year grants will run out in 2011, and when that happens it could cause a nasty shock. Barring a new windfall—and none is in sight—NIH's budget will drop sharply next year. Much of the work recently begun will be left short of cash. The result could be the lowest grant funding rates in NIH history, and the academic job market will suddenly dry up—especially for young researchers.

"There will be a reduction in capacity. How abruptly and exactly when it hits we're still trying to piece together," says Howard Garrison, public affairs director for the Federation of American Societies for Experimental Biology (FASEB) in Bethesda, Maryland. Garrison issued a gloomy overview of the figures last month.

Last week, in his first budget defense as NIH director before a House of Representatives Appropriations subcommittee, Francis Collins was asked about the poststimulus problem. He showed a graph of NIH's budget over the years with a label at the point where the stimulus money runs out and the total will drop \$4 billion in 2011, assuming NIH gets the \$32.2 billion President Barack Obama has requested (see graph). (NIH's baseline budget, not counting stimulus money, would increase by 3.2%.)

"This is the cliff that people are talking about," Collins said. "We are going to face a crunch" in 2011, he said. The success rate, or portion of reviewed applications that receive funding, which hovered around 30% a few years ago and 20% this year, "will be more like 15%." The committee chair, Representative David Obey (D-WI), asked whether the NIH windfall was a mistake. No, Collins replied, it has been a "wonderful investment."

The looming cliff may not look scary to some NIHers because they have weathered booms and busts before. NIH's budget

doubled between 1999 and 2003, encouraging academic research institutions to go on a building spree, bringing new researchers into the system, and spurring others to expand their labs with multiple grants. After that, NIH's budget stayed essentially flat, or fell if adjusted for inflation, until 2009–10. Then the stimulus money in the American Recovery and Reinvestment Act (ARRA) threw a lifeline to labs that were in jeopardy; many NIH panels reached below the "payline" and funded grants that had just missed the cutoff for funding from NIH's regular budget.

The reprieve probably won't last: In its own projection, NIH says it may fund just 35,202 grants in 2011, a drop of 4377. But the exact size of that drop will depend on

tions could swell further if a significant fraction of the 19,000 or so scientists whose ideas were rejected for a special ARRA competition for high-impact research submit those ideas as regular R01 grant applications.

The number of applications will also depend on how many investigators with ARRA funding request a 1-year unfunded extension, which would stretch their 2-year grants to 3 years. NIH says it will be "gracious" about granting such requests, to help smooth the poststimulus transition, Collins testified. In addition, certain institutes, such as the National Cancer Institute, have pledged to fund many ARRA grants for another 2 years.

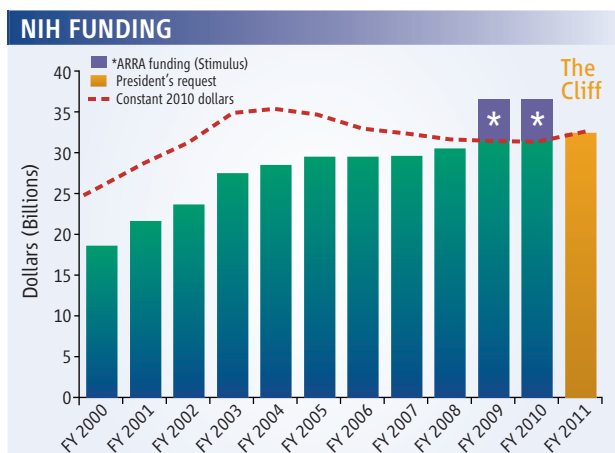
The cliff problem was the focus of a meeting of economists, academic leaders, and NIH officials last week at Cold Spring Harbor Laboratory on Long Island in New York. Economist Paula Stephan of Georgia State University in Atlanta expects trouble as a glut of postdoctoral fellows and graduate students who were hired or kept on with ARRA money will hit the job market. "There is going to be a huge backlog of individuals looking for jobs and fellowships in 2011 and 2012," predicts Stephan.

One suggestion: persuade Congress to fund one-time "bridge" fellowships to give ARRA-funded postdocs more time to find academic jobs or switch careers. "It wouldn't take a whole lot of money," says meeting organizer Richard Freeman, a labor economist at Harvard University. Looking beyond 2011, the larger challenge is to find ways for NIH to smooth the impact of roller-coaster budgets, meeting participants agreed.

The grim outlook may not have hit university labs just yet. Stem cell researcher Michael Kyba of the University of Minnesota Medical School in Minneapolis, who won three ARRA grants, says he's been too busy trying to get papers out to start applying for the next grants to continue these studies—although he should be. "Grants are built on papers," he explains. "I will worry about [applying for] grants in a few months."

Harvard microbiologist and American Society for Microbiology President Roberto Kolter, who spoke at Cold Spring Harbor, says "there's a wait-and-see attitude right now" in his department, which didn't apply for much ARRA money. He and many of his colleagues won't need to renew their grants for 2 or 3 years. But then, he expects, "there will be some casualties."

—JOCELYN KAISER



Hard landing. An abrupt drop in NIH's budget in 2011 could mean a plunge in grant success rates and a glut of postdocs on the job market.

several factors. The first is NIH's final budget, which usually ends up larger than the president's request. Various groups are pushing for more than Obama's \$32 billion: Research!America says the agency needs \$35 billion, a 13.5% increase; FASEB suggests \$37 billion; and Senator Arlen Specter (D-PA), who engineered NIH's \$10.4 billion in total ARRA funding, says the agency should get \$40 billion. But whether NIH will come anywhere close to that is unclear in a still-tight fiscal climate.

Another uncertainty is how many research grant applications NIH will receive in 2011; these make up the denominator of the success rate figure. After reaching an all-time high of about 47,500 in 2007, the number of proposals was declining or at least leveling off, NIH data suggest. The numbers could rise again, however, if investigators with the one-time ARRA grants apply for new awards. Applica-

PLANETARY SCIENCE

Mission to Probe Venus's Curious Winds And Test Solar Sail for Propulsion

TOKYO—Venus is considered Earth's twin because of its similar size and composition. But Venus's carbon dioxide-laden atmosphere and thick sulfuric clouds are nothing like Earth's. And most puzzlingly, venusian winds move at up to 60 times the speed of the planet's rotation; Earth's fastest winds clock in at just 10% to 20% of the rotation speed.

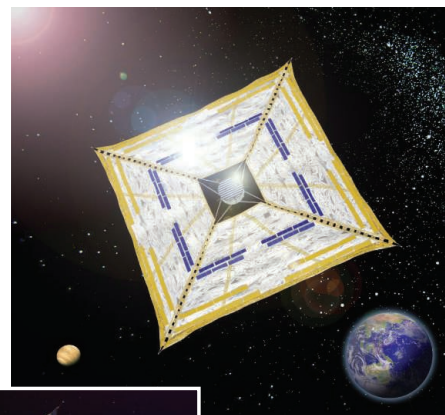
The force behind this superrotation is "one of the most outstanding problems in planetary geofluid dynamics," says Kevin Baines, a planetary scientist at NASA's Jet Propulsion Laboratory in Pasadena, California, and the University of Wisconsin, Madison. The Japan Aerospace Exploration Agency (JAXA) is out to solve this problem with its Akatsuki mission, slated for launch on 18 May. Akatsuki will train four cameras at ultraviolet and infrared wavelengths on the planet to track clouds at different altitudes. The craft also has a high-speed camera that aims to capture venusian lightning, presumed to occur but never directly observed.

The superfast winds might result from thermal waves generated in the solar-heated clouds or from a complex interaction of atmospheric circulation patterns, says

Takeshi Imamura, a planetary scientist at JAXA's Institute of Space and Astronautical Science in Sagami-hara. "It is also possible that our observations [will eliminate] both of these theories and find new mysteries," he says.

Akatsuki, to be placed in an equatorial orbit, will complement the European Space Agency's Venus Express, which has been in a polar orbit around the planet since 2006. The two teams are planning coordinated observations using different instruments. Baines, a member of the Venus Express science team, expects the findings to provide clues to understanding the evolution of Earth's atmosphere and the effects of greenhouse gases. "There is a lot Venus can teach us" about Earth, he says.

In addition to sending Akatsuki on its way, JAXA's H-IIA rocket is carrying what will become, if all goes well, the first spacecraft that uses the pressure of photons from the sun for propulsion. Two previous attempts foundered on rocket failures. The Interplanetary Kite-craft Accelerated by Radiation Of



Sail away. An innovative ribless design will use centrifugal force to unfurl the solar sail of the IKAROS spacecraft (inset).

the Sun (IKAROS) mission will use a square polyimide sail 20 meters across diagonally and a mere 0.0075 mm thick. "The most important point is the deployment," says Jun'ichiro Kawaguchi, a JAXA scientist who conceived IKAROS. Other designs have used struts to keep the sail taut. IKAROS relies on centrifugal force to spin out its sail. If the mission succeeds, Kawaguchi says, a solar sail could be paired with ion engines, used in JAXA's Hayabusa mission (*Science*, 30 April, p. 565), to speed spacecraft to the solar system's outer reaches.

—DENNIS NORMILE

ScienceNOW

From *Science's* Online Daily News Site

Bald Eagle Comeback Could Come at Expense of Other Species

Fishing, farming, and pesticides have decimated the bald eagle population of California's Channel Islands. But current efforts to reintroduce the birds could likewise threaten some of the islands' other endangered species, including foxes, according to a new study. <http://bit.ly/baldeagle-comeback>



Scientists Resurrect Mammoth Hemoglobin

By inserting a 43,000-year-old woolly mammoth gene into *Escherichia coli* bacteria, scientists have figured out how these ancient beasts adapted to the subzero temperatures of prehistoric Siberia and North America. The gene, which codes for the oxygen-transporting protein hemoglobin, allowed the animals to keep their tissues supplied with oxygen even at very low temperatures. <http://bit.ly/mammoth-hemoglobin>

New Madrid Quakes May Not Have Been So Colossal

In the winter of 1811–12, four major earthquakes shook what is today southeastern Missouri. The first of the temblors flattened the small town of New Madrid and, according to some eyewitness accounts, temporarily reversed the course of the Mississippi River. Other reports over the years have claimed that the quakes could be felt as far away as the U.S. East Coast, where they rattled furniture in Washington, D.C., and rang church bells in Charleston, South

Carolina. Together, the New Madrid quakes have become known as one of the most powerful natural events in U.S. history. But were they? <http://bit.ly/newmadridquakes>

Island Reptiles Buck an Evolutionary Trend

The Caribbean island of Martinique has a lot to offer: beautiful beaches, tropical weather, and a perfect place to challenge notions about how new species arise. When anole lizards arrived on Martinique more than 8 million years ago, there were four separate, smaller islands. Over time, this physical isolation should have allowed the lizards to evolve into different species, according to conventional evolutionary wisdom. But that's not what happened. <http://bit.ly/lizard-evolution>



Read the full postings, comments, and more at news.sciencemag.org/sciencenow.

U.S. GRADUATE EDUCATION

Report Seeks Expansion in a Time of Belt-Tightening

Timing is everything. And for a new report that urges the U.S. government to spend billions more on graduate education while scolding universities for not doing enough to attract and mentor the students who would earn those degrees, the timing is unfortunate.

*The Path Forward: The Future of Graduate Education in the United States** is a thoughtful and data-rich analysis of how graduate education can increase productivity and foster innovation by raising the quality of the U.S. workforce. Its authors, an 18-member panel convened by the Council of Graduate Schools and the Educational Testing Service, regard the report as an essential follow-up to the improbably influential 2005 report from the National Academies, *Rising Above the Gathering Storm*, which successfully made the case for hiking federal spending on research and education.

The new report urges graduate departments to shorten the amount of time it takes to earn a degree, reduce attrition rates, recruit and retain more Hispanic and African-American students, and provide better career counseling. In return, it says, the federal government should safeguard access to graduate education by international students and expand support for master's and doctoral programs. Toward that last goal, the report proposes a new cross-agency doctoral training initiative that would finance the educations of an additional 25,000 students annually at a cost of \$10 billion over 5 years.

But while lawmakers are likely to applaud the call for reforming graduate education, the idea of a massive boost in fed-

eral support flies in the face of current fiscal realities. The prospect of a \$1.3 trillion budget deficit and a freeze next year on most domestic programs may cause policymakers to put this report on their bookshelves rather than in their legislative agendas.

"There is a lot of interest in this topic within the Obama Administration," says Under Secretary Martha Kanter, who oversees postsecondary education at the Department of Education and who spoke at the report's unveiling last week. "But there are always a lot of competing ideas. And most of them don't get to the finish line. That's especially true in a flat budget year, which it looks like 2011 will be. What we don't do as much of is ask, 'What can we stop doing?'"

By coincidence, *The Path Forward* was released the day after the House of Representatives Science and Technology Committee approved a reauthorization of the 2007 law, the America COMPETES Act. That legislation drew heavily from the academies' *Gathering Storm* report and provided the justification for Congress to boost funding for research. But the committee's actions on 28 April indicate how even the staunchest supporters of academic research and training have had to curb their enthusiasm.

During an 8-hour markup of the bill (H.R. 5116), the committee remained solidly in favor of the Administration's promise to double over 10 years the budgets of the National Science Foundation (NSF), the Department of Energy's (DOE's) research programs, and the National Institute of Standards and Technology. In addition to spelling out operating procedures for DOE's new energy hubs and its Advanced Research Projects Agency-Energy,

creating an Office of Innovation within the Department of Commerce, and giving NSF the authority to tackle knotty research challenges with prizes rather than grants, the bill also embraces several new federal initiatives to improve science and math education.

At the same time, however, the committee bowed to growing congressional concern about the federal deficit by lopping 10% off authorized agency spending levels spelled out in a version of the bill introduced only a few days earlier by the committee's chair, Representative Bart Gordon (D-TN). The lower levels were proposed by Gordon himself, who had anticipated the push for a lower number by crafting an amendment to his own bill that pared its original cost from \$93 billion to \$82 billion over 5 years. The amendment passed with bipartisan support.

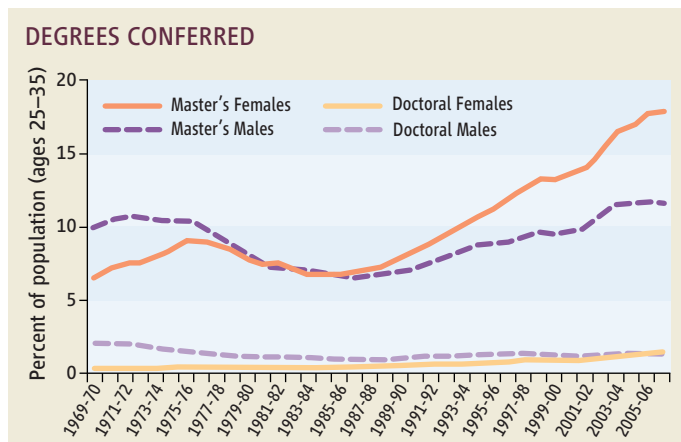
"These [new] levels are lower than I'd like them to be," Gordon acknowledged at the start of the session. "But I consider them to be practical. ... We will maintain a doubling path but on a slightly less steep trajectory."

Authorizing a new doctoral training program would require the science committee to find new sources of revenue or cut into existing programs. Neither is a "winnable argument" under the current fiscal restraints, says one committee aide.

Gordon, who is not running for reelection in November, hopes the full House will pass the COMPETES bill before the end of May. Its reauthorization would be a crowning achievement of a 26-year congressional career. But will the Senate go along? Although the bill has many supporters, the Senate's legislative calendar is already jammed.

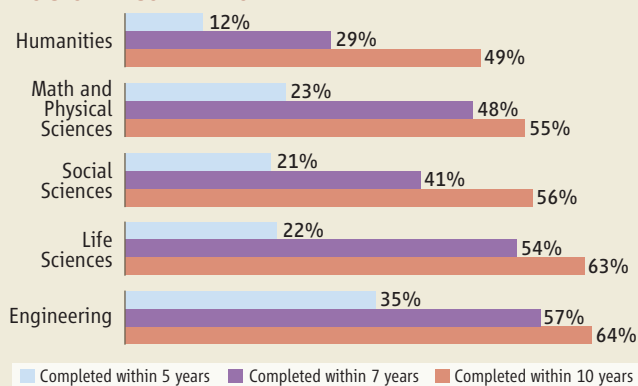
—JEFFREY MERVIS

*http://www.fgereport.org/rsc/pdf/CFGE_report.pdf



Gender difference. Women account for most of the growth in the number of U.S. graduate degrees awarded since 1995.

DOCTORAL COMPLETION RATE



Getting it done. Engineers finish their Ph.D.s faster than those in other disciplines. But there's a significant dropout rate in every field.

LAB SAFETY

Radiation Accident a 'Wake-Up Call' For India's Scientific Community

NEW DELHI—The improper disposal of a derelict gamma-ray research device at the University of Delhi has resulted in the death of a scrap-metal worker—and drawn scrutiny of how India's academic institutions handle radioactive materials.

India's nuclear watchdog, the Atomic Energy Regulatory Board (AERB) in Mumbai, has labeled the incident as a "seri-

a campuswide spring cleaning to create space for newly recruited staff. On 26 February, the university auctioned to a scrap dealer a Gammacell 220 research irradiator, which university chemist B. K. Sharma had imported from Atomic Energy of Canada Ltd. in 1968 but which had lain unused since Sharma's retirement in 1985. At the scrap

yard, unaware that the machine they were dismantling contained cobalt-60, seven workers fell ill; two are in critical condition. One died—India's first known fatality from a radiation accident. Some of the cobalt-60, a gray-blue metal resembling nickel, is unaccounted for.

The incident has raised questions about the University of Delhi's competence in radiation safety. An auction committee comprised of top university chemists had determined that radiation levels from the Gammacell's cobalt-60 would be "manageable," Pental says, as the isotope has a half-life of 5.27 years. (In fact, the material in such a device can release

lethal radiation for decades.) AERB's chair, nuclear scientist S. S. Bajaj, told *Science* that he was "shocked" by the Delhi researchers' "ignorance." Bajaj acknowledges that the fate of the Gammacell device was not being tracked by AERB and should never have "disappeared from [AERB's] radar," but he says the university was obliged to inform AERB before disposing of the device. An agency team would have helped ensure that this was done safely. AERB has launched an audit of the roughly 140 Gammacells in India's academic institutions.

Pental says he takes "moral responsibility" for any negligence on the part of the University of Delhi. AERB and Delhi police are now conducting an investigation; anyone found guilty of unauthorized disposal of a radioactive source can be jailed for up to 5 years. "Zero tolerance," says Parthasarathy, "is the only way forward."

—PALLAVA BAGLA



In the hot seat. University of Delhi's Deepak Pental says he takes moral responsibility for the botched disposal of a cobalt-60 source (inset).

ous violation" and slapped India's premier university with an indefinite moratorium on radiation-related activities. That "such an outstanding university was so callous is mind-blowing," declares former AERB Secretary K. S. Parthasarathy, who says the tragedy "should be a wake-up call." On 30 April, India education minister Kapil Sibal ordered universities to beef up guidelines for the handling and use of hazardous materials.

India's accident is the latest in a series of radioactive mishaps worldwide. Since 1993, the International Atomic Energy Agency in Vienna has logged several dozen incidents involving the loss or theft of "dangerous" radioactive sources, including cobalt-60. One of the worst cases of cobalt-60 exposure occurred at a scrap yard in Samut Prakash, Thailand, in February 2000, when the dismantlement of a radiation-therapy unit killed three people and injured 10.

The accident in India occurred after plant biotechnologist Deepak Pental, the University of Delhi's vice chancellor, ordered

Science Insider

From the *Science* Policy Blog



The nation's **first cancer vaccine** will soon hit the market. The U.S. Food and Drug Administration has approved for sale Provenge, meant to treat prostate cancer that has spread throughout the body. Patients given the vaccine lived about 4 months longer, nearly 26 months total, than controls. A course of treatment costs \$93,000. <http://bit.ly/91nzld>

The National Institutes of Health has approved **four stem cell lines** submitted by WiCell, the nonprofit associated with the University of Wisconsin, Madison. The lines were approved under the Bush-era stem cell policy, but they had to undergo new scrutiny to make sure they meet stiff ethics rules. <http://bit.ly/cfmzR6>

The plight of postdocs, and the glacial pace of negotiations between a union that represents them in California and the University of California (UC), was the subject of a fairly acrimonious congressional hearing. About 6000 postdocs—10% of all U.S. postdocs—work at UC. The union is seeking a pay structure modeled on federal guidelines. <http://bit.ly/a5GqXi>

A new handbook offers hints for universities trying to increase **diversity** while staying within the law. The joint effort of AAAS (which publishes *Science*) and the Association of American Universities includes tips, examples of what works, and legal analyses. <http://bit.ly/bfIM8p>

Drug company scientists released news of a **high-profile clinical trial** of a drug to treat fragile X syndrome by divulging the results to a reporter at *The New York Times*. Researchers feel "pretty good about the [unpublished] data," said a scientist with Novartis. <http://bit.ly/9CPJHh>

Are **genomewide association studies** useful? A provocative paper in *Cell* says no, arguing that by linking common gene variants to increased disease risks, they fail to uncover helpful connections. A better approach, says the much-discussed paper, may be to look for rare variants. <http://bit.ly/bA3MOQ>

See the full postings and more at news.sciencemag.org/scienceinsider.



Close Encounters Of the Prehistoric Kind

The long-awaited sequence of the Neandertal genome suggests that modern humans and Neandertals interbred tens of thousands of years ago, perhaps in the Middle East

IT'S THE MYSTERY OF MOUNT CARMEL. ON this limestone ridge overlooking the coast of Israel, modern humans lived in caves off and on for tens of thousands of years, starting more than 100,000 years ago. Then, perhaps as early as 80,000 years ago, members of another species reached and occupied the caves: heavy-bodied Neandertals, who were escaping a cold spell in Europe and moving south into the Middle East. Did the two species meet here? Did they mate?

The archaeological record in the caves is ambiguous on that question, and anthropologists have fought bitterly over it. Some claim that the anatomy of fossils shows that Neandertals, our closest cousins, did mate with modern humans, either in the Middle East or in Europe. But others thought modern humans coming out of Africa completely replaced Neandertals with little or no interbreeding. And the genetic evidence from ancient bones showed no sign that Neandertals had swapped genes with our ancestors—until now.

On page 710, an international team of researchers presents their first detailed analysis of the draft sequence of the Neandertal genome, which now includes more than 3 billion nucleotides collected from the bones of three female Neandertals who lived in Croatia more than 38,000 years ago. By comparing this composite Neandertal genome with the complete genomes of five living humans from different parts of the world, the researchers found that both Europeans and Asians share 1% to 4% of their nuclear DNA with Neandertals. But Africans do not. This suggests that early modern humans interbred with Neandertals after moderns left Africa, but before they spread into Asia and Europe. The evidence showing interbreeding is “incontrovertible,” says paleoanthropologist John Hawks of the University of Wisconsin, Madison, who was not involved in the work. “There’s no other way you can explain this.”

Kissing cousins. A few Neandertals mated with early modern humans and passed on some of their genes to living humans.

As a result, many people living outside Africa have inherited a small but significant amount of DNA from these extinct humans. “In a sense, the Neandertals are then not altogether extinct,” says lead author Svante Pääbo, a paleogeneticist at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, who was surprised to find he was part Neandertal. “They live on in some of us.”

The team also used the Neandertal DNA like a probe to find the genes that make us modern. Even though the genomes of humans and Neandertals are 99.84% identical, the researchers identified regions that have changed or evolved since our ancestors and Neandertals diverged sometime between 270,000 and 440,000 years ago—their new, slightly younger estimate of the split. So far, the team has detected tantalizing differences in genes involved in metabolism, skin, the skeleton, and the development of cognition, although no one knows yet how these genetic changes affect physiology. “This is a groundbreaking study!” enthuses evolutionary geneticist Hendrik Poinar of McMaster University in Hamilton, Canada. “We can actually discuss an extinct human species—Neandertals—on a genetic level rather than strictly on morphological grounds.”

Mixed marriage

The discovery of interbreeding in the nuclear genome surprised the team members. Neandertals did coexist with modern humans in Europe from 30,000 to 45,000 years ago, and perhaps in the Middle East as early as 80,000 years ago (see map, p. 681). But there was no sign of admixture in the complete Neandertal mitochondrial (mtDNA) genome or in earlier studies

of other gene lineages (*Science*, 13 February 2009, p. 866). And many researchers had decided that there was no interbreeding that led to viable offspring. “We started with a very strong bias against mixture,” says co-author David Reich of Harvard Medical

School in Boston. Indeed, when Pääbo first learned that the Neandertal DNA tended to be more similar to European DNA than to African DNA, he thought, “Ah, it’s probably just a statistical fluke.” When the link persisted, he thought it was a bias in the data. So the researchers used different methods in different labs to confirm the result. “I feel

Online
sciencemag.org
S Special online feature and podcast on the Neandertal genome.

confident now because three different ways of analyzing the data all come to this conclusion of admixture,” says Pääbo.

The finding of interbreeding refutes the narrowest form of a long-standing model that predicts that all living humans can trace their ancestry back to a small African population that expanded and completely replaced archaic human species without any interbreeding. “It’s not a pure Out-of-Africa replacement model—2% interbreeding is not trivial,” says paleoanthropologist Chris Stringer of the Natural History Museum in London, one of the chief architects of a similar model. But it’s not wholesale mixing, either: “This isn’t like trading wives from cave to cave; the amount of admixture is tiny,” says molecular anthropologist Todd Disotell of New York University in New York City. “It’s replacement with leakage.”

Although the 1.3-fold coverage of the Neandertal genome is a remarkable technical feat, one-third of the genome is still murky. In a separate paper (p. 723), the team describes and successfully tests a new method for filling in gaps in the rough draft of the genome.

The team also used three methods to nail down the interbreeding result. First, they compiled the Neandertal genome using DNA from the limb bones of three female Neandertals who lived in Vindija Cave in Croatia from 38,000 to 44,000 years ago; they confirmed parts of the genome with much smaller amounts of DNA from Neandertals who lived in Spain, Germany, and Russia.

Once they were satisfied that the composite genome was a fair representation of Neandertals from across a great part of their geographical range, researchers compared the Neandertal genome to a chimpanzee’s to determine which genetic variants were primitive, ancestral forms. Then they compared the new, derived genetic variants in Neandertals to those in the complete genomes of five living humans, including a San from Southern Africa, a Yoruba from West Africa, a Papua New Guinean, one Han Chinese, and one French European.

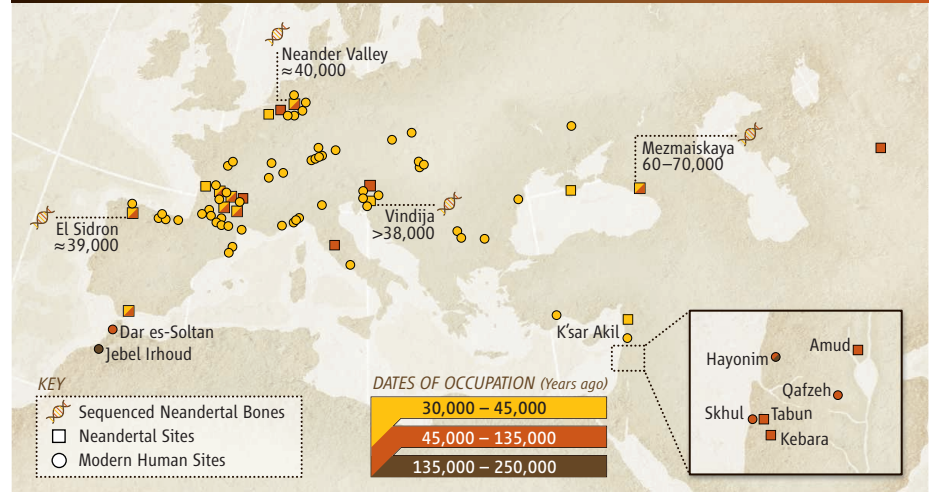
The team measured the genetic proximity of Neandertals to pairs of modern humans from different continents, first using single-nucleotide polymorphisms (SNPs), or sites in the genome where a single nucleotide differs between individuals. When they compared a Neandertal with a European and an Asian, they found that the Neandertal always shared the same amount of derived (or more recently evolved) SNPs with each of them. But when they compared a Neandertal with an African and a European, or with an African and an Asian, the Neandertal always shared



Clean cave. Researchers carefully worked with DNA from three Neandertals’ bones (left) in Vindija Cave, Croatia, where the fossils were found (right).



NEANDERTAL AND MODERN HUMAN OVERLAP



Points of contact. Archaeological data suggest that Neandertals and early modern humans may have overlapped early in the Middle East and later in Europe.

more SNPs with the European or Asian than with the African. “We’ve shown that Neandertals are significantly more closely related to non-Africans than Africans on average,” says Reich.

Even though they looked at just two Africans for this part of the study, those two have a particularly ancient, diverse heritage, so they are a good proxy for much of the genetic diversity in Africa. But sequencing additional Africans would be a good idea, says Reich.

For now, it seems Neandertals interbred with the ancestors of Europeans and Asians, but not with the ancestors of Africans. At first, “we were baffled that this affinity with

Neandertals was not only in Europe and West Asia [where it was most expected], but also in Papua New Guinea” where Neandertals never set foot, says Pääbo.

To be certain, they used two other methods to detect gene flow between Neandertals and Eurasians. Using the published genome of an African American from the Human Genome Project, they compared large regions of African and European ancestry in this single genome to Neandertal regions. In this person’s genome, the European and Neandertal segments were more similar to each other than either was to the African segments.



Cloned Neandertals Still in the Realm of Sci-Fi

Science-fiction writers have been resurrecting Neandertals in novels for decades, imagining what it would be like to see and communicate (not to mention mate) with another species of human. So once the idea of sequencing the Neandertal genome became more than a glimmer in a paleogeneticist's eye, some have asked, "Could we, should we, would we, bring this extinct human species back to life?" After all, biologists are trying to bring back the woolly mammoth by cloning. But for both technical and ethical reasons, experts say, bringing back a Neandertal is a pipe dream.

Could we do it? Robert Lanza laughed at the thought. Chief scientific officer for Advanced Cell Technology in Worcester, Massachusetts, he and his colleagues have cloned species from

cows to goats to mice and extended their efforts to include endangered species and human embryos. But cloning Neandertals is fantasy, says Lanza. "You can't clone from stone, and you can't clone from DNA that has been destroyed from weather and the elements," he points out.

The Neandertal genome sequence reported on page 710 (and see main text, p. 680) reflects the battered state of the starting DNA, which came from bones that are 38,000 to 44,000 years old. Because the isolated DNA was in pieces typically about 50 bases long, there are many missing stretches, particularly repetitive regions. "We will never have a finished sequence for the Neandertal the way we have for a mouse," says Svante Pääbo, who led the Neandertal sequencing project at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. *Jurassic Park* aside, recon-

structing an organism with a partial genome would be like constructing a building with a partial blueprint.

Even if scientists had the complete genome, it wouldn't be enough. DNA itself doesn't tell the whole story. Chemical modifications to the genome, the way chromosomes arrange in the nucleus, and maternal components in the egg all play a role in translating a genetic blueprint into a viable individual. "It's not just the DNA; there's a lot else going on," says Lanza. None of that information is even available for Neandertals.

Then, too, cloning doesn't typically start with a genome; it starts with two cells. One cell provides a nucleus (with DNA inside), and one is an egg cell, most often of the same species, whose DNA has been removed. The nucleus is then transferred to the egg, sometimes by fusing the two cells. "If you have just got DNA, you are asking an enormous amount of the oocyte that you are going to put the DNA into," explains Ian Wilmut, who cloned Dolly the sheep and now works at the University of Edinburgh in the United Kingdom. "It has to reform the nucleus and reprogram [the DNA]."

That leads to the next problem: What species' egg would play host to this DNA? The obvious candidate would be a modern human egg, but they are notoriously fickle and don't take well to nuclear transfer, even of modern human DNA. "There's something different about primates that we haven't identified," says Wilmut. "[Cloning] works very poorly." And incompatibilities between Neandertal DNA and the human egg might further diminish the chances of a viable embryo.

Molecular geneticist George Church of Harvard University has proposed another approach: modify the DNA in a human cell line to resemble the Neandertal. "This is a daunting task, but with future technological developments and enough time and money, it may be possible," says Adrian Briggs, who worked on the

Finally, population geneticist Rasmus Nielsen of the University of California (UC), Berkeley, scanned the human genome for "ancient" genomic segments—those that might predate the time when modern humans arose, about 200,000 years ago. Before receiving the Neandertal DNA sequences, he identified 13 genomic regions that were unusually variable, and therefore likely to be evolutionarily ancient, in 48 people outside of Africa. He identified 13 "old" variants as possibly coming from Neandertals or other archaic ancestors, because they were missing from the genomes of 23 African Americans (used as proxies for Africans). Then the team looked

in the Neandertal genome—and found 10 of the 13 ancient variants. "There are places in the genome where we can say this section is really, really likely to be from a Neandertal," says Reich.

When and where did modern humans pick up those Neandertal genes? The most likely scenario "was the movement of a few Neandertals into a group of moderns," says co-author and population geneticist Montgomery Slatkin of UC Berkeley. If a few Neandertals interbred with members of a small population of modern humans, Neandertal gene variants might persist in subsequent generations of modern humans if the interbred population expanded rapidly,

thereby spreading Neandertal DNA widely.

This scenario apparently fits with fossils and stone tool data from the Israeli caves such as Skhul, Qafzeh, and Tabun, where Neandertals show up in the region as early as 80,000 years ago, when moderns were already there. Although each group may have occupied the caves intermittently, some say they may have overlapped for up to 10,000 years. Neandertals and moderns apparently even occupied the same cave, Tabun, at different times. The two species had much in common: Both lived in caves, used similar toolkits (although Neandertals may have made better spear points), and hunted the same fallow deer and gazelles.

Neandertal genome sequence and is about to join Church's lab. In theory, one could convert a human or chimp genome to a Neandertal genome—base by base—while it is still nicely nestled in a stem cell, then clone it. But there's on the order of a million differences between the Neandertal and human genomes, and the more changes needed, the greater the risk of introducing errors.

If, somehow, a viable embryo were produced, this developing chimera would need a surrogate mother. What species would that mother belong to? Again, the obvious choice is a human, but no one knows whether a modern woman's biochemistry would be compatible with that of a Neandertal fetus. And is it ethical for a human surrogate mother to birth a Neandertal baby? Church thinks ethical views will evolve as technology improves. Once cloning works well in a variety of animals and stem cell-derived organs become commonplace, "I think the resistance to it will disappear," he says.

But others disagree. "We do not—and should not—create human beings just to satisfy our scientific curiosity," says Pääbo, pointing out that Neandertals are a species of human, so cloning them raises many of the same ethical issues as cloning a modern human.

Cloning Neandertals would involve several "ethically deplorable steps," says Briggs, including using surrogate mothers and risking high failure rates, abnormal births, and, sometimes, early death of clones. With a Neandertal, "all of these safety issues would apply, only writ large," says Wilmut. And how would a Neandertal fit into modern human society? "I see no palatable conditions," says Pääbo. "Not even for medical purposes are we thinking about creating a [modern] human being. Why would we consider something like this, which is much less pressing?"

—E.P.

"It doesn't surprise me," says archaeologist Ofer Bar-Yosef of Harvard University about the ancient DNA finding. "We always predicted low-level mixing," because some Neandertals in the Middle East, such as a female skeleton at Tabun, look less robust than Neandertals in Asia and Europe. Mixing in this region could also have happened later, when another group of modern humans came out of Africa about 60,000 years ago and perhaps met Neandertals, who were still occupying caves in the Middle East until 50,000 years ago, says Stringer.

Finally, the researchers cannot rule out the possibility that what they see as "Neandertal" motifs are really ancient genetic

Computer Kid Makes Good

Late 2007 was a real low point for Richard "Ed" Green and colleagues at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. A year earlier, in *Nature*, they had predicted that they could sequence the Neandertal genome using 20 grams of bone and 6000 runs using "next generation" sequencing technologies. They knew going in that most of the DNA in fossil bone is bacterial, with only a small percentage of Neandertal DNA. But it turned out that the bones to be sequenced had far less Neandertal DNA than the sample on which they based their projections. "We were in kind of an awkward situation of having announced to the world we were going to do it, and we were left with no concrete plan of how to do it," Green recalls. "That was very scary." Their fears increased when they discovered that their first million bases of Neandertal sequence were contaminated with modern human DNA.

"But we worked it out along the way," says Green, the postdoctoral fellow in charge of the project. He and colleagues developed methods to control contamination by putting bar codes on all DNA coming from the fossils (*Science*, 13 February 2009, p. 866). They cut down on the amount of DNA to be deciphered by cutting up much of the bacterial DNA so that the sequencing reactions ignored it. Everyone, especially Green, stresses the team effort involved. "Many people here have been able to say they 'saved the Neandertal genome project,'" he notes. And yet Green, 37, still stands out.

"Ed brought the quantitative and algorithmic horsepower needed to interpret the Neandertal data," says David Haussler of the University of California (UC), Santa Cruz, where Green now works as an assistant professor. "He invented new analysis methods that allowed the Neandertal project to happen."

That computational horsepower is what landed Green the job of shepherding the Neandertal genome. After getting a degree in computational biology from UC Berkeley, he joined Svante Pääbo's lab at the Max Planck institute in 2005 to explore

Problem solver. Richard Green overcame obstacles to sequencing Neandertal DNA.

the evolution of genes that can code for more than one protein. Pääbo and the sequencing company 454 Life Sciences in Branford, Connecticut, had just sequenced cave bear and mammoth DNA and were puzzling over the results: There was so much microbial sequence, it was hard to detect mammalian DNA. Green knew what to do: He enlisted a cluster of computers to compare the DNA with that of known sequences, including dog and elephant, so he could discard the microbial sequence and focus on the tiny bit of mammalian DNA.

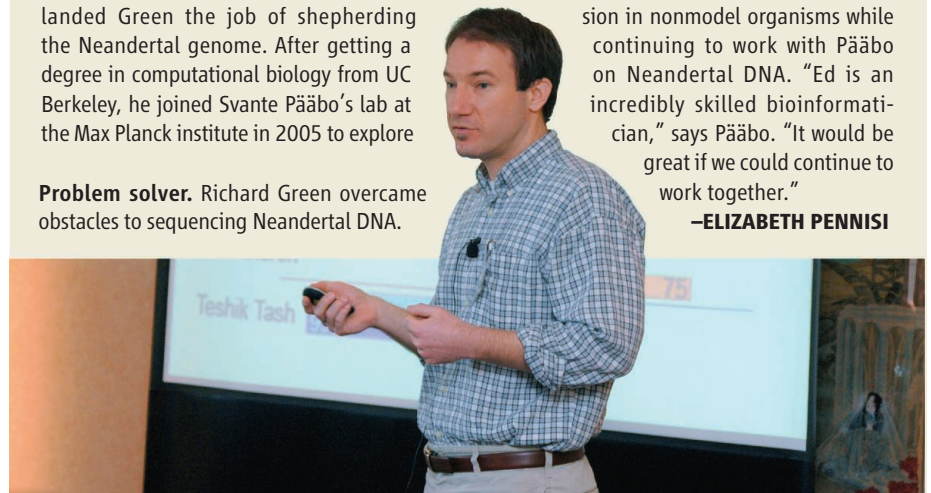
"This was really the first large-scale snapshot of what the universe of [ancient] DNA looked like when it came out of a bone," Green recalls. "Then Svante said, 'Let's try Neandertal.' It was obvious that this was a once-in-a-lifetime opportunity."

He took charge of the bioinformatics effort, writing software to better detect Neandertal DNA and to deal with degradation. "He is able to design ways to analyze a whole genome under circumstances that are nonstandard," says Pääbo.

Green also coordinated the design and logistics of the rest of the project, which involved about 50 people. He was "very patient in terms of helping and training others," says Pääbo. Former graduate student Adrian Briggs agrees: "Without Ed's enthusiasm and competence, the project would never have proceeded so fast." The job required long hours at the lab, but Green says he didn't mind because the Max Planck facilities were "maximally comfortable," complete with Ping-Pong table, sauna, barbecue grill, and even a resting room.

Switching gears was not new to Green, who had started off in developmental biology as an undergrad and studied cancer biology in grad school before moving to computational biology. Now that he's settled in at UC Santa Cruz, Green expects to switch gears again. He wants to look at gene expression in nonmodel organisms while continuing to work with Pääbo on Neandertal DNA. "Ed is an incredibly skilled bioinformatician," says Pääbo. "It would be great if we could continue to work together."

—ELIZABETH PENNISI





SEPARATING THEM FROM US

Some genes that differ between modern humans and Neandertals

Gene	Significance
RPTN	Encodes the protein repetin, expressed in skin, sweat glands, hair roots, and tongue papilli
TRPM1	Encodes melastatin, a protein that helps maintain skin pigmentation
THADA	Associated with type 2 diabetes in humans; evolutionary changes may have affected energy metabolism
DYRK1A	Found in an area critical for causing Down syndrome
NRG3	Mutations associated with schizophrenia
CADPS2, AUTS2	Mutations implicated in autism
RUNX2 (CBRA1)	Causes cleidocranial dysplasia, characterized by delayed closure of cranial sutures, malformed clavicles, bell-shaped rib cage, and dental abnormalities
SPAG17	Protein important for the beating of the sperm flagellum

variants that Neandertals and some modern humans inherited from a common ancestor they shared before Neandertals split off. Although all early modern populations, including in Africa, interbred, that gene flow was not complete enough to pass these Neandertal motifs to all Africans. Human populations that were more closely related to the ancestors of Neandertals carry those motifs while Africans do not, says Reich.

To date, the genomic data don't support interbreeding in the time and place when everyone most expected it: between 45,000 and about 30,000 years ago in Europe. Neandertals and moderns lived in such proximity in France, for example, that some researchers think Neandertals imitated modern stone-tool and beadmaking technologies. But such late European mixing cannot explain the current findings, in which Asians and Europeans are equally similar to Neandertals. It's still possible that Neandertals and modern humans in Europe interbred rarely and that the Neandertal genes were swamped out in a large population of modern humans, says Slatkin.

In some ways, it is surprising that there isn't more evidence of interbreeding, now that researchers know it was biologically possible. "For some reason, they didn't interbreed a lot—something was preventing them," says evolutionary geneticist Sarah Tishkoff of the University of Pennsylvania. "Was it a cultural barrier?"

Modern motifs

The Neandertal genome also gives researchers a powerful new tool to fish for genes that have evolved recently in our lineage, after we split from Neandertals. The team compared the Neandertal genome with the genomes of five diverse modern humans. They found 78

Different paths. A partial list of genes that differ between Neandertals (*left*, reconstruction from Amud Cave, Israel) and early modern humans (*right*, reconstruction from Qafzeh Cave, Israel).

new nucleotide substitutions that change the protein-coding capacity of genes and that are present in most humans today; just five genes had more than one such substitution. That's a tiny fraction of the 3 billion bases in each genome. "Only 78 substitutions in the last 300,000 years!" says Poinar. "The fact that so few changes have become fixed on the human lineage is amazing."

But the mutations they've found so far "are all very interesting, precisely because there are so few," says Pääbo, whose team is trying to identify their function. The catalog includes changes in genes that encode proteins important for wound healing, the beating of sperm flagellum, and gene transcription (see table, above). Several of these newly evolved modern human genes encode proteins expressed in the skin, sweat glands, and inner sheaths of hair roots, as well as skin pigmentation. "The fact that three of six genes carrying multiple substitutions are in skin is fascinating," says Poinar. Pääbo speculates that these changes "reflect that skin physiology has changed but how, of course, we don't know yet."

Some of those changes are likely to be neutral changes that accumulated through genetic drift, but the team also used the Neandertal data to find other evolutionary changes that were beneficial to modern humans and so rose to high frequencies in some populations. Specifically, they have identified 15 regions containing between one and 12 genes. The widest region is located on chromosome 2 and contains the gene *THADA*, a region that varies in modern humans and that has been associated

with type 2 diabetes. Changes in this gene may have affected energy metabolism in modern humans.

Other mutations appear to be in genes important in cognitive development and that, when mutated in living people, contribute to diseases such as Down syndrome, schizophrenia, and autism. One gene, *RUNX2*, is associated with a disease that leads a spectrum of developmental abnormalities, including misshapen clavicles and a bell-shaped rib cage. Suggestively, Neandertals had bell-shaped rib cages and possibly peculiar clavicles. But precisely how all these genetic differences are expressed physiologically is the next frontier. "We need to follow up. Are there regions that are functionally significant?" says Tishkoff. By 7 May, the Neandertal data should be posted on Ensembl and the UC Santa Cruz browser, so other teams can do just that, says Pääbo.

His own group is already working on such functional studies. Postdoctoral researcher Matthias Gralle is analyzing the way these recently evolved genetic differences change the way proteins are expressed. Such studies may eventually offer clues about why Neandertals went extinct—and our ancestors didn't. "The mystery isn't just why they disappeared," says paleoanthropologist Jean-Jacques Hublin of the Max Planck Institute for Evolutionary Anthropology. "It is why we were so successful that we replaced all the others." For now, researchers are delighted that this "groundbreaking" genomic work has made it possible to ask such interesting questions, says Poinar. "This is the real appeal of this project: What will the genome of the Neandertal tell us about functional differences between the two [species]," says Poinar.

—ANN GIBBONS

PROFILE: FRÉDÉRIQUE DARRAGON

Unraveling a Riddle in Plain Sight

Amateur archaeologist Frédérique Darragon has spent 12 years documenting hundreds of mysterious towers in Southwest China—and winning over skeptical academics

CHENGZI, CHINA—The jeep grinds over a rise on a rutted dirt road in the foothills of the Himalayan Mountains. “Stop here!” exclaims Martine Francoise Darragon. The svelte socialite-turned-explorer leaps out and holds up an old photo showing snow-capped peaks towering over gentle valley slopes, with a rock-strewn river in the foreground. The black-and-white landscape of western Sichuan Province, captured in the 1930s by the intrepid botanist Joseph Rock, matches the view from where Darragon, who goes by the name Frédérique, is standing. “This is where he took the picture,” she says in her French–New York City accent. But something from the photo is missing from the land today: two eye-catching stone towers whose beveled walls, viewed from above, would form eight-pointed stars.

In the 70 years since Rock’s travels through the Tribal Corridor of Tibet and western Sichuan, a small village inhabited by Minyag people has sprung up here on the banks of the Chengzi River, some 3750 meters above sea level. Defying the thin air, Darragon, 60, bounds across a stone bridge over the Chengzi and makes a beeline to where the nearest tower in the postcard once stood. All that’s left now of a structure that had been at least 25 meters tall is a dilapidated first-floor section: interior stonework and fill, and a doorway framing blue sky. Over more than a decade, Darragon has identified nearly

1000 such ancient structures in

Sichuan and neighboring Tibet, from total wrecks like this one to largely intact towers exceeding 50 meters in height. Many more have been lost through the ages.

Why these Himalayan towers were built is an enduring mystery. Other structures—squat, square towers erected against northern invaders—once were widespread and are of scant scientific interest. But the much taller star-shaped towers and other arresting buildings suggest that the medieval kingdoms of these lands were more ingenious and sophisticated than many scholars have presumed. Clusters of towers on mountain slopes may have been status symbols in a game of one-upmanship among wealthy merchants, Darragon says, while lone towers in river valleys likely served as lookouts or way stations on the southwestern Silk Road. Some towers may have held religious meaning. And a star-shaped design may help withstand shaking in a region prone to strong earthquakes.

Over the past decade, Darragon has had fragments of wood beams from several dozen towers radiocarbon-dated, yielding approximate ages ranging from 300 to 1700 years old. Most presumably were built during this period, although any single tower’s age is hard to pin down: Some beams could have been replaced after a tower was built, yielding a more recent carbon-14 date, while others could have been built using beams from older trees that predated the towers.

Darragon may be

an amateur, but her sleuthing and derring-do have earned the respect of Chinese scientists. “Some experts did not know what to think of her at first,” says Zhong Xiao-Hou, director of the National Architecture Institute of China in Beijing. “But we have come to admire her spirit and enthusiasm for our heritage.”

Drawing on Darragon’s work, the State Administration of Cultural Heritage of China (SACH) is expected to soon nominate dozens of the more imposing structures to UNESCO’s World Heritage list as the Diaolou Buildings and Villages of Tibetan and Qiang Ethnic Groups Cultural Landscapes. The towers “represent an extraordinary heritage and tradition, and deserve to be fully preserved,” says Francesco Bandarin, assistant director-general for culture at UNESCO.

But a UNESCO listing may not come in time for a clutch of towers near Danba, in western Sichuan, that are imperiled by the construction of a hydropower dam. The rGyal-rong towers have become Darragon’s latest cause célèbre.



Star attraction. Western Sichuan’s star-shaped Bamei tower, restoration of which was completed last year, likely was built in the 13th or 14th century.

Unlikely savior

Darragon's wanderings in the highlands of Sichuan and Tibet, a forbidding swath of land she has come to know better than most Westerners—and Chinese—are a far cry from the racy days of her youth. As a teenager from a wealthy Parisian family, Darragon spent summers riding horses in England and winter breaks skiing in the Swiss Alps. All the while, she nurtured a counterculture side: "I was a communist at heart," says Darragon, who worked one summer on a kibbutz in Israel.

She was also a playgirl at heart. When Darragon was 18, she inherited a small fortune from her father, an inventor and machinemaker. In early 1971, she sailed across the Atlantic as a bikini-clad deckhand in the first Cape Town-to-Rio de Janeiro race, then from Brazil to the West Indies before returning to Paris just in time to take final exams and graduate from the Université Paris X de Sciences Economiques.

Darragon spent a few years managing her real estate and modeling. Then in the summer of 1978, she took up polo and was the first woman to play at the Bagatelle Polo Club in Paris before moving to Buenos Aires and becoming a record-setting player in Argentina. If those pursuits were not enough, Darragon has also raced as a jockey on thoroughbred horses, won renown as a samba dancer in Rio, and has had a lifelong passion for oil painting.

Along the way, Darragon collected prominent boyfriends and near-death experiences. The former include cable TV mogul Ted Turner, whom she has known since 1969 and lived with from 2000 to 2003. The latter include wiping out during a high-speed motorcycle chase in which she smashed her Suzuki into a car, flew over the hood, and landed 10 meters away without a scratch, and taking a polo ball in the mouth that crushed her jaw and knocked out several teeth. (She wrapped a scarf around her head and finished the game.) "At times I do regret my happy-go-lucky life," says Darragon, who has no children and never married. But then she found a higher purpose in China.

Darragon visited China for the first time in the early 1990s and afterward spent months each year backpacking across the country. In 1993, her interest in the endangered snow leopard brought her to Tibet, where she would have her most serious brush with death. On a solo trek in the Himalayas near the border with Bhutan in 1996, Darragon decided to shelter overnight in a tiny cave. It was freezing, so she lit a fire inside. "Terrible idea," says Darragon, whose Chinese name, Bing



A higher purpose. Frédérique Darragon samples wood for radiocarbon analysis. World Heritage status would be a boost for Minyang and other minorities who tend the towers.

Yan, means "ice flame." The fire sucked the scant oxygen from the thin air and, Darragon says, "I felt a snap inside my head." She had suffered a stroke and couldn't control her left side, but managed to drag herself outside the cave before blacking out. Three days after she recovered consciousness, Tibetan shepherds came across her and carried her to a village, where she caught a ride to Lhasa, Tibet's capital. Enfeebled, Darragon holed up in the Holiday Inn—at the time, the only hotel in Lhasa with room service. "I could only crawl and didn't want my mother to see me in such a state," she jokes. Four months later, she felt well enough to fly back to France.

Darragon was soon back in Tibet and venturing where few foreigners had ever gone, such as the remote valleys of

Gongbu Jiangda, which once belonged to the ancient Nyangpo Kingdom. It was in places like that, off the beaten track, where Darragon encountered the stone towers, including ones with the astonishing star-shaped walls. Locals did not know who had built the towers, how old they were, or why they were built.

Archaeologists knew about some of the more accessible towers, and Taoping, a Qiang village with several towers a few hours from Chengdu, Sichuan's capital, has long been a tourist draw. Darragon has spent a total of 5 years roaming the hinterlands of Sichuan and Tibet, analyzing more than 250 standing towers, including a few dozen star-shaped ones and 750 or so other ruins. "The sheer amount of data she has collected about these architectural curiosities will be welcomed by a variety of specialists," says John Vincent Bellezza, senior research fellow at the Tibet Center of the University of Virginia in Charlottesville.

Under the auspices of the Unicorn Foundation, a U.S. nonprofit that Darragon founded in 2001 with seed money from Turner, wood samples from 77 towers—54 in Sichuan and 23 in Tibet—have been radiocarbon-dated by Beta Analytic in Miami, Florida. With Darragon's help, over the past 6 years Achim Bräuning, a dendrochronologist at the University of Erlangen-Nürnberg in Germany, has been building a tree-ring database from the region. He has radiocarbon-dated wood from 16 towers; preliminary results, he says, corroborate Darragon's data. The earliest date—318 C.E., with a margin of error of 40 years—is from a star-shaped tower in Nyangpo. "It's probably the oldest one still standing in the world," Darragon says.

Search for meaning

Solving the riddle of the towers is a daunting challenge. When Darragon queried locals, she often just got shrugs. Sometimes it was lack of knowledge; sometimes it was a communication barrier. The region is a Tower of Babel of mutually unintelligible languages, and Mandarin Chinese won't get you far. Chinese annals from the Han Dynasty, which lasted from 206 B.C.E. to 220 C.E., refer to tall towers, according to Chen Zongxiang, a retired historian in Chengdu. Some old towers along the Min River in Sichuan are spaced several kilometers apart on a clear line of sight and must have served as watchtowers or beacons, says Yasuhiko Nagano, an expert on the Tibetan rGyalrong languages at the National Museum of Ethnology in Osaka, Japan. Untold numbers of smaller tow-

ers were built during the Jinchuan Wars of the 18th century, when the Manchurian emperor Qian Long sought to pacify the region. But many of the star-shaped towers lack classic features—arrow slits, for example—of defensive fortifications.

Scholars believe that the towers served various purposes. A majestic, hill-hugging assemblage in Danba may have arisen as merchants vied to outdo each other with taller and taller structures, similar to the origin of the San Gimignano towers in Italy, says Darragon. Other towers may have symbolized the *dmu* cord that in Tibetan lore connects heaven and Earth. In rGyalrong lands, now part of western Sichuan, “it appears that some towers were built to propitiate the deities” by ceremonially projecting ritual participants into a sky realm, says Bellezza.

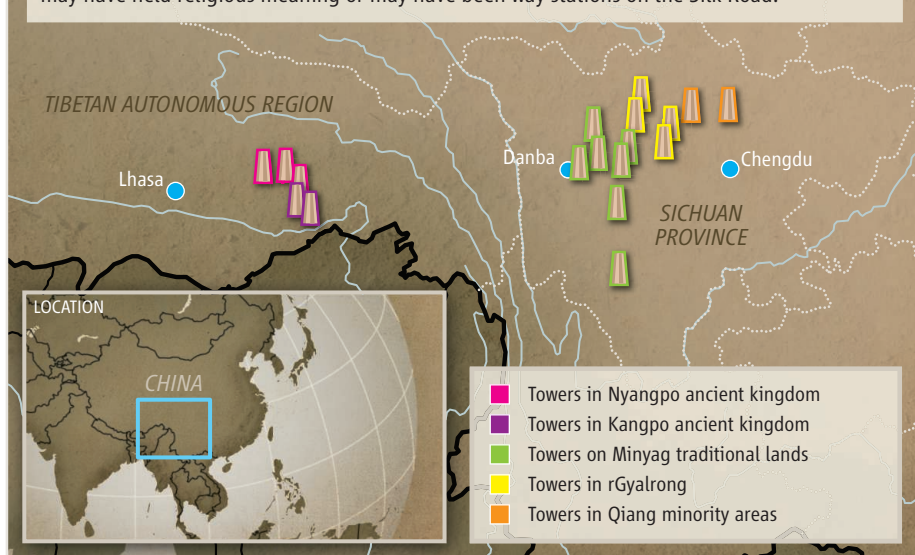
One idea of Darragon’s gaining support is that towers were way stations along the southern Silk Road, which passed through Tibet and Sichuan. “The more fabulous towers all lie along the trade routes,” she says. Many are located in strategic spots in river valleys along routes traveled by medieval caravans. According to Zhong, these towers “may have embodied the economic strength of each village.” Major commodities in the region then were silk, salt, tea, and musk of the forest musk deer. “Beyond the fact that we know that this area was exporting valuable musk in medieval times, we don’t know very much about the mechanics of the musk trade,” says Anya King, a historian at the University of Southern Indiana in Evansville. Chinese scholars have postulated a “Musk Road” linking with the Silk Road in the Ngari region of western Tibet. “It is entirely possible that the towers served as regional depots of the trade routes, but at this point it won’t be easy to prove,” King says. To test the idea, Darragon is attempting to organize a dig to sample soil for silk, tea, and musk remains at one undisturbed medieval tower.

The star-shaped construction, with its distinctive five to 13 points, is found almost nowhere else. Only a handful of such towers exist outside the region, in India, Iran, Tajikistan, and Afghanistan—including the Bahram Shah Minaret in Ghazni. The shape may help resist earthquakes. “That’s what local people tell us,” says Li Chunxia, an anthropologist at Sichuan University in Chengdu who has collaborated with Darragon.

Earthquakes are not the main threat, however. Impoverished villagers have blown up or dismantled towers for building materials, Li says. An especially grievous loss was two ancient Nyangpo towers destroyed in 2006

TOWERS OF THE HIMALAYAS

Cultural landscape. Some towers in Sichuan and Tibet were defensive fortifications. Others may have held religious meaning or may have been way stations on the Silk Road.



and 2008, Darragon says. She has come across three villages named Bajiaodiao, or “eight-angle fortress,” only to find nothing but scattered stones. “The remaining towers desperately need protection,” she says. Tourism could help. The Unicorn Foundation has assisted some communities to restore towers and convert farmhouses into guesthouses. “If rightly harnessed, some of the proceeds could go to scientific pursuits in the region,” says Bellezza.

Another big boost for the preservation effort, Darragon says, would be to include the towers on UNESCO’s World Heritage List. Over the years, Darragon and Unicorn have submitted maps, radiocarbon data, and oral history to SACH in support of a nomination. “She is so diligent. Even though she is not a professional, her work is hard to refute,” Li says. “Frédérique is a force,” adds Bandarin. “She has identified the value of this heritage, she has promoted the necessary research and scientific investigation, and she has attracted the interest of the local and national governments and of UNESCO,” he says. SACH is expected to decide soon on whether to nominate several groups of towers as Cultural Landscape sites. “The cultural self-esteem of about 50,000 minority people living in these regions is riding on it,” Darragon says.

If the nomination comes through, Darragon says she can rest easy—and move on. “I’ve never spent such a long time on any one thing in my life,” she says. But Darragon has set herself one last tower-related task: saving several rGyalrong tow-

ers near Danba, which have special significance to her. She glimpsed towers in China for the first time on a trip to Danba, when in 1997, she recalls, “in the pouring rain I caught sight of tall structures clinging to far-away mountain slopes.”

Several of these towers, including the tallest in Sichuan that is about 700 years old, are in jeopardy. “They will be under water,” Darragon says, if a second dam planned for the area is built. Chinese colleagues and other experts have joined with her to lobby authorities to revise the hydropower plans.

A recent trip to Danba allowed Darragon to reinforce that message. In Badi village high in the Danba hills, three young girls dash out of their home, smiling and giggling, as Darragon’s jeep pulls up. She and her entourage are invited into the rGyalrong home for bowls of homemade cottage cheese and butter tea. They’ve known Darragon for years and treat her like family.

After a chat, Darragon takes her leave and gets back to work. As the afternoon shadows lengthen, she climbs into a ruined section of a fortress with an attached tower and uses a penknife to gouge a chunk from a wooden beam that she’ll send for radiocarbon analysis. “That’s it for today,” she says. Tomorrow she has more survey work and a meeting with Danba officials. “I will do my best to persuade them that a World Heritage nomination would benefit Danba more than a second dam would,” she says. Given her tenacity, it would be a wonder if her latest quest were to fail.

—RICHARD STONE

Evidence
ignored

692

Building trust in
climate science

695

Supplying tellurium
for photovoltaics

699



LETTERS | BOOKS | POLICY FORUM | EDUCATION FORUM | PERSPECTIVES

LETTERS

edited by Jennifer Sills

Climate Change and the Integrity of Science

WE ARE DEEPLY DISTURBED BY THE RECENT ESCALATION OF POLITICAL ASSAULTS ON SCIENTISTS in general and on climate scientists in particular. All citizens should understand some basic scientific facts. There is always some uncertainty associated with scientific conclusions; science never absolutely proves anything. When someone says that society should wait until scientists are absolutely certain before taking any action, it is the same as saying society should never take action. For a problem as potentially catastrophic as climate change, taking no action poses a dangerous risk for our planet.

Scientific conclusions derive from an understanding of basic laws supported by laboratory experiments, observations of nature, and mathematical and computer modeling. Like all human beings, scientists make mistakes, but the scientific process is designed to find and correct them. This process is inherently adversarial—scientists build reputations and gain recognition not only for supporting conventional wisdom, but even more so for demonstrating that the scientific consensus is wrong and that there is a better explanation. That's what Galileo, Pasteur, Darwin, and Einstein did. But when some conclusions have been thoroughly and deeply tested, ques-

tioned, and examined, they gain the status of “well-established theories” and are often spoken of as “facts.”

For instance, there is compelling scientific evidence that our planet is about 4.5 billion years old (the theory of the origin of Earth), that our universe was born from a single event about 14 billion years ago (the Big Bang theory), and that today's organisms evolved from ones living in the past (the theory of evolution). Even as these



are overwhelmingly accepted by the scientific community, fame still awaits anyone who could show these theories to be wrong. Climate change now falls into this category: There is compelling, comprehensive, and consistent objective evidence that humans are changing the climate in ways that threaten our societies and the ecosystems on which we depend.

Many recent assaults on climate science and, more disturbingly, on climate scientists by climate change deniers are typically driven by special interests or dogma, not by an honest effort to provide an alternative theory that credibly satisfies the evidence. The Intergovernmental Panel on Climate Change (IPCC) and other scientific assessments of climate change, which involve thousands of scientists producing massive and comprehensive reports, have, quite expectedly and normally, made some mistakes. When errors are pointed out, they are corrected. But there

is nothing remotely identified in the recent events that changes the fundamental conclusions about climate change:

(i) The planet is warming due to increased concentrations of heat-trapping gases in our atmosphere. A snowy winter in Washington does not alter this fact.

(ii) Most of the increase in the concentration of these gases over the last century is due to human activities, especially the burning of fossil fuels and deforestation.

(iii) Natural causes always play a role in changing Earth's climate, but are now being overwhelmed by human-induced changes.

(iv) Warming the planet will cause many other climatic patterns to change at speeds unprecedented in modern times, including increasing rates of sea-level rise and alterations in the hydrologic cycle. Rising concentrations of carbon dioxide are making the oceans more acidic.

(v) The combination of these complex climate changes threatens coastal communities and cities, our food and water supplies, marine and freshwater ecosystems, forests, high mountain environments, and far more.

Much more can be, and has been, said by the world's scientific societies, national academies, and individuals, but these conclusions should be enough to indicate why scientists are concerned about what future generations will face from business-as-usual practices. We urge our policy-makers and the public to move forward immediately to address the causes of climate change, including the unrestrained burning of fossil fuels.

We also call for an end to McCarthy-like threats of criminal prosecution against our colleagues based on innuendo and guilt by association, the harassment of scientists by politicians seeking distractions to avoid taking action, and the outright lies being spread about them. Society has two choices: We can ignore the science and hide our heads in the sand and hope we are lucky, or we can act in the public interest to reduce the threat of global climate change quickly and substantively. The good news is that smart and

effective actions are possible. But delay must not be an option.

P. H. GLEICK,* R. M. ADAMS, R. M. AMASINO, E. ANDERS, D. J. ANDERSON, W. W. ANDERSON, L. E. ANSELIN, M. K. ARROYO, B. ASFAW, F. J. AYALA, A. BAX, A. J. BEBBINGTON, G. BELL, M. V. L. BENNETT, J. L. BENNETZEN, M. R. BERENBAUM, O. B. BERLIN, P. J. BJORKMAN, E. BLACKBURN, J. E. BLAMONT, M. R. BOTCHAN, J. S. BOYER, E. A. BOYLE, D. BRANTON, S. P. BRIGGS, W. R. BRIGGS, W. J. BRILL, R. J. BRITTEN, W. S. BROECKER, J. H. BROWN, P. O. BROWN, A. T. BRUNGER, J. CAIRNS JR., D. E. CANFIELD, S. R. CARPENTER, J. C. CARRINGTON, A. R. CASHMORE, J. C. CASTILLA, A. CAZENAVE, F. S. CHAPIN III, A. J. CIECHANOVER, D. E. CLAPHAM, W. C. CLARK, R. N. CLAYTON, M. D. COE, E. M. CONWELL, E. B. COWLING, R. M. COWLING, C. S. COX, R. B. CROTEAU, D. M. CROTHERS, P. J. CRUTZEN, G. C. DAILY, G. B. DALRYMPLE, J. L. DANGL, S. A. DARST, D. R. DAVIES, M. B. DAVIS, P. V. DE CAMILLI, C. DEAN, R. S. DEFRIES, J. DEISENHOFER, D. P. DELMER, E. F. DELONG, D. J. DEROSIER, T. O. DIENER, R. DIRZO, J. E. DIXON, M. J. DONOGHUE, R. F. DOOLITTLE, T. DUNNE, P. R. EHRLICH, S. N. EISENSTADT, T. EISNER, K. A. EMANUEL, S. W. ENGLANDER, W. G. ERNST, P. G. FALKOWSKI, G. FEHER, J. A. FERREJOHN, A. FERSHT, E. H. FISCHER, R. FISCHER, K. V. FLANNERY, J. FRANK, P. A. FREY, I. FRIDOVICH, C. FRIEDEN, D. J. FUTUYMA, W. R. GARDNER, C. J. R. GARRETT, W. GILBERT, R. B. GOLDBERG, W. H. GOODENOUGH, C. S. GOODMAN, M. GOODMAN, P. GREENGARD, S. HAKE, G. HAMMEL, S. HANSON, S. C. HARRISON, S. R. HART, D. L. HARTL, R. HASELKORN, K. HAWKES, J. M. HAYES, B. HILLE, T. HÖKFELT, J. S. HOUSE, M. HOUT, D. M. HUNTEN, I. A. IZQUIERDO, A. T. JAGENDORF, D. H. JANZEN, R. JEANLOZ, C. S. JENCKS, W. A. JURY, H. R. KABACK, T. KAILATH, P. KAY, S. A. KAY, D. KENNEDY, A. KERR, R. C. KESSLER, G. S. KHUSH, S. W. KIEFFER, P. V. KIRCH, K. KIRK, M. G. KIVELSON, J. P. KLINMAN, A. KLUG, L. KNOPOFF, H. KORNBORG, J. E. KUTZBACH, J. C. LAGARIAS, K. LAMBECK, A. LANDY, C. H. LANGMUIR, B. A. LARKINS, X. T. LE PICHON, R. E. LENSKE, E. B. LEOPOLD, S. A. LEVIN, M. LEVITT, G. E. LIKENS, J. LIPPINCOTT-SCHWARTZ, L. LORAND, C. O. LOVEJOY, M. LYNCH, A. L. MABOGUNJE, T. F. MALONE, S. MANABE, J. MARCUS, D. S. MASSEY, J. C. MCWILLIAMS, E. MEDINA, H. J. MELOSH, D. J. MELTZER, C. D. MICHENER, E. L. MILES, H. A. MOONEY, P. B. MOORE, F. M. M. MOREL, E. S. MOSLEY-THOMPSON, B. MOSS, W. H. MUNK, N. MYERS, G. B. NAIR, J. NATHANS, E. W. NESTER, R. A. NICOLL, R. P. NOVICK, J. F. O'CONNELL, P. E. OLSEN, N. D. OPDYKE, G. F. OSTER, E. OSTROM, N. R. PACE, R. T. PAINE, R. D. PALMITER, J. PEDLOSKY, G. A. PETSCH, G. H. PETTINGILL, S. G. PHILANDER, D. R. PIPERNO, T. D. POLLARD, P. B. PRICE JR., P. A. REICHARD, B. F. RESKIN, R. E. RICKLEFS, R. L. RIVEST, J. D. ROBERTS, A. K. ROMNEY, M. G. ROSSMANN, D. W. RUSSELL, W. J. RUTTER, J. A. SABLOFF, R. Z. SAGDEEV, M. D. SAHLINS, A. SALMOND, J. R. SANES,

R. SCHEKMAN, J. SCHELLNHUBER, D. W. SCHINDLER, J. SCHMITT, S. H. SCHNEIDER, V. L. SCHRAMM, R. R. SEDEROFF, C. J. SHATZ, F. SHERMAN, R. L. SIDMAN, K. SIEH, E. L. SIMONS, B. H. SINGER, M. F. SINGER, B. SKYRMS, N. H. SLEEP, B. D. SMITH, S. H. SNYDER, R. R. SOKAL, C. S. SPENCER, T. A. STEITZ, K. B. STRIER, T. C. SÜDHOF, S. S. TAYLOR, J. TERBORGH, D. H. THOMAS, L. G. THOMPSON, R. T. TJIAN, M. G. TURNER, S. UYEDA, J. W. VALENTINE, J. S. VALENTINE, J. L. VAN ETEN, K. E. VAN HOLDE, M. VAUGHAN, S. VERBA, P. H. VON HIPPEL, D. B. WAKE, A. WALKER, J. E. WALKER, E. B. WATSON, P. J. WATSON, D. WEIGEL, S. R. WESSLER, M. J. WEST-EBERHARD, T. D. WHITE, W. J. WILSON, R. V. WOLFENDEN, J. A. WOOD, G. M. WOODWELL, H. E. WRIGHT JR., C. WU, C. WUNSCH, M. L. ZOBACK

*To whom correspondence should be addressed. E-mail: petergleick@pacinst.org

Notes

1. The signatories are all members of the U.S. National Academy of Sciences but are not speaking on its behalf.
2. Signatory affiliations are available as supporting material at www.sciencemag.org/cgi/content/full/328/5979/689/DC1.

Shifting the Debate on Geoengineering

AS DISCUSSED IN THE RECENT POLICY FORUM “The politics of geoengineering” (J. J. Blackstock and J. C. S. Long, 29 January, p. 527), there is growing recognition that avoiding dangerous climate change during the 21st century may require society to adopt geoengineering technologies to supplement CO₂ emission reduction efforts. Unfortunately, despite the essential role

that CO₂ removal (CDR) and solar radiation management (SRM) technologies may play in reducing the risks of dangerous climate change, discussions of the necessary research and development [including the Policy Forum and others (1, 2)] frequently turn into debates about the environmental costs and benefits of SRM. A more productive approach would shift the debate to comparing the relative costs and benefits of CDR and SRM.

CDR approaches are frequently discounted because, as Blackstock and Long explain, “technical challenges and large uncertainties [surround] large-scale CDR deployment.” Although this may be true for human-built systems that capture CO₂ from air at ambient concentrations, there are other technologies based on biological carbon fixation that could be fast-tracked for rapid deployment during the next few decades (3). Most major international energy corporations are investing in algal-based biofuel technologies because of the tremendous production potential of algae relative to terrestrial energy crops (4). Commercial-scale production of algal biofuels will begin during the next 5 years, and rapid scaling up can be expected afterward if the economic incentives are favorable. However, becoming carbon negative will require society to develop plans for retrofitting existing coal-fired power plants and building future ones so that they can burn algal biomass and capture the emitted CO₂ for subsequent sequestration. The basic technologies described here are not novel; rather, I am proposing a conceptual rearrangement that may enable society to transition more gracefully

CORRECTIONS AND CLARIFICATIONS

Research Articles: “Doc2b is a high-affinity Ca²⁺ sensor for spontaneous neurotransmitter release” by A. J. Groffen *et al.* (26 March, p. 1614). Several author affiliations were not footnoted properly; three corrected affiliations follow. Y. Takai, Department of Biochemistry and Molecular Biology, Kobe University Graduate School of Medicine, Kobe 650-0017, Japan. J. G. Borst, Department of Neuroscience, Erasmus MC, University Medical Center, Rotterdam, 3000 CA, Netherlands. N. Brose, Max-Planck-Institut für Experimentelle Medizin, Abteilung Molekulare Neurobiologie, 37075 Göttingen, Germany.

Letters: “Oil and water do mix” by J. L. Kavanau (19 February, p. 958). Due to an editorial error, the title was incorrect. It should have been “Opposites attract.”

Reports: “100-million-year dynasty of giant planktivorous bony fishes in the Mesozoic seas” by M. Friedman *et al.* (19 February, p. 990). The author Matt Friedman’s affiliation should have been “Committee on Evolutionary Biology, University of Chicago, 1025 East 57th Street, Chicago, IL 60637, USA.” The affiliation that was listed is his present address.

News of the Week: “DSM-V at a glance” by G. Miller and C. Holden (12 February, p. 770). In the sidebar, it was reported that the term “gender identity disorder” has been retained. In fact, a different term—“gender incongruence”—has been proposed.

Research Articles: “PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice” by F. Baudat *et al.* (12 February, p. 836). M. Lichten was incorrectly listed as an author in references 18 and 19. The correct authors for reference 18 are C. Grey, F. Baudat, and B. de Massy; for reference 19, the correct authors are E. D. Parvanov, S. H. Ng, P. M. Petkov, and K. Paigen.

Reports: “Epigenetic transgenerational actions of endocrine disruptors and male fertility” by M. D. Anway *et al.* (3 June 2005, p. 1466). As clarification of the abstract to Anway *et al.*, the F₁ to F₄ generations were examined after vinclozolin treatment, and F₁ and F₂ generations were examined after methoxychlor treatment. To clarify data referred to in the last paragraph of the Report, serum testosterone measurements after vinclozolin treatment were shown in reference 21 (Uzumcu *et al.*) for the F₁ generation. Data for the F₁ to F₄ generations were subsequently published in Anway *et al.*, *J. Androl.* **27**, 868 (2006). Serum testosterone measurements after methoxychlor treatment were shown in reference 20 (Cupp *et al.*) for the F₁ generation, but measurements of the F₂ generation have not been published. The *Science* Anway *et al.* manuscript showed DNA methylation analysis after vinclozolin treatment, but the DNA methylation data after methoxychlor treatment have not been published.

from fossil to modern carbon fuel sources while simultaneously reducing CO₂ levels in the atmosphere and ocean.

CHARLES H. GREENE

Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY 14853, USA. E-mail: chg2@cornell.edu

References

1. A. Robock *et al.*, *Science* **327**, 530 (2010).
2. D. W. Keith, E. Parson, M. G. Morgan, *Nature* **463**, 426 (2010).
3. D. W. Keith, M. Ha-Duong, J. K. Stollaroff, *Climat. Change* **74**, 17 (2006).
4. M. E. Huntley, D. G. Redalje, *Mitigation Adapt. Strategies Global Change* **12**, 573 (2007).

Response

GREENE SUGGESTS THAT CO₂ REMOVAL methods deserve expanded evaluation and research. We agree. In the long run, these methods may be the only way to reduce atmospheric concentrations of CO₂ to values closer to those of the preindustrial era. Greene suggests a scheme for using biomass to generate electricity combined with carbon capture and storage. This idea has merit. Even schemes that capture CO₂ directly from the air deserve expanded research.

However, Greene's statement that "discussions of the necessary research and

development...frequently turn into debates about the environmental costs and benefits of SRM [solar radiation management]" misses a key point motivating all three of the articles he cites [our Policy Forum and (1, 2)]. The two approaches differ in both strategic impact and risks. Most CO₂ removal schemes, including those suggested by Greene, would be slow acting and expensive, and would pose no transboundary risks. In contrast, SRM techniques appear inexpensive and could have rapid climatic impact, but present a host of global climatic and political risks.

The low cost and technical feasibility of some SRM technologies (particularly stratospheric aerosol injection) mean that SRM might be our only response if a "climate emergency" develops. However, these traits also mean that SRM could be globally tested unilaterally by a single country, to the possible detriment of others (3). Beyond the climatic risks this presents, such actions could also severely disrupt progress on international climate policy.

The discussion of urgent governance challenges in the articles Greene cites is not a distraction; it is central to figuring out how

to safely and prudently conduct research into SRM technologies. No such acute research governance challenges exist for most CO₂ removal techniques.

JASON J. BLACKSTOCK^{1,2*} AND JANE C. S. LONG³

¹International Institute for Applied Systems Analysis, Laxenburg A2361, Austria. ²Centre for International Governance Innovation, Waterloo, ON N2L 6C2, Canada. ³Lawrence Livermore National Laboratory, Livermore, CA 94550, USA.

*To whom correspondence should be addressed. E-mail: jjb@iiasa.ac.at

References

1. A. Robock *et al.*, *Science* **327**, 530 (2010).
2. D. W. Keith, E. Parson, M. G. Morgan, *Nature* **463**, 426 (2010).
3. D. G. Victor, M. G. Morgan, J. Apt, J. Steinbruner, *Foreign Aff.* **88**, 64 (2009).

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 3 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

EVOLUTION

Two Critics Without a Clue

Douglas J. Futuyma

What Darwin got wrong, according to Jerry Fodor and Massimo Piattelli-Palmarini, was natural selection. They will fell evolutionary ethics; social, epistemological, and psychological Darwinism; and all such ills by cutting “the tree at its roots.” The authors accept the historical and entirely material reality of evolution and are forthright atheists who want no part of creationism. But, like many earlier critics, they cannot accept that natural selection is the mechanism of adaptive evolution.

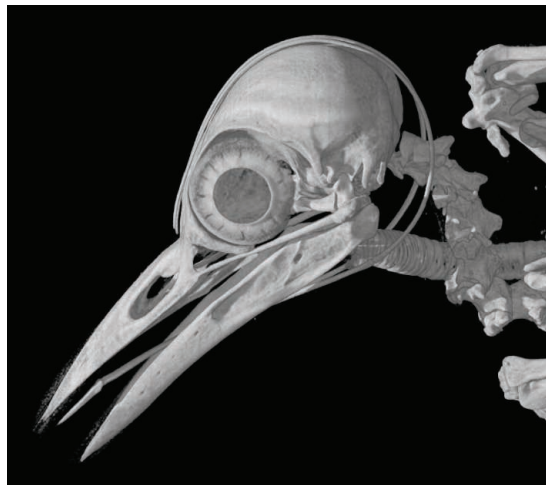
Fodor (a philosopher and cognitive scientist at Rutgers University) and Piattelli-Palmarini (a cognitive scientist at the University of Arizona) attempt to make their case by biological and philosophical arguments. Here, I address their claims from the viewpoint of a biologist, as I do not pretend to much knowledge of other fields. The authors equate evolutionary theory (or neo-Darwinism) with an “adaptationist” view that natural selection is virtually omnipotent—a view today held by hardly any knowledgeable evolutionary biologists. Their neo-Darwinian straw man claims that “random,” unlimited variation originates by mutation. In contrast, they say, newly arisen variation is highly nonrandom because it is channeled by internal (e.g., developmental) constraints: pigs lack wings because their evolved anatomy precludes heritable phenotypic variation from which selection could possibly forge wings. Such constraints, they hold, largely explain convergent evolution and phenotypes generally. Some phenotypes also originate, they claim, by self-organization, as a direct consequence of physical and chemical laws. None of these arguments is new, and none is a substitute for natural selection.

Evolutionary biologists have long understood that mutation is random with respect to context-specific advantage, not with respect to the kinds of variations that can arise. Phylogenetically local constraints, a major subject of research today, have been recognized by “neo-Darwinians” such as Ernst Mayr:

What Darwin Got Wrong

by Jerry Fodor and Massimo Piattelli-Palmarini

Farrar, Straus and Giroux, New York, 2010. 286 pp. \$26. ISBN 9780374288792. Profile, London. 280 pp. £20. ISBN 9781846682193.



Of endogenous origin? Skull of golden-fronted woodpecker, *Melanerpes aurifrons*.

Earlier authors ... had a far greater interest in the phenotypic potential bestowed on an evolutionary line by its epigenetic heritage than recent evolutionists. This must be kept in mind when we speak of the randomness of mutations.... [T]he epigenotype sets severe limits to the phenotypic expression of such mutations; it restricts the phenotypic potential. The understanding of this limitation facilitates the understanding of evolutionary parallelism and polyphyletic evolution. (1)

But Fodor and Piattelli-Palmarini take endogenous constraint much further than almost any biologists do. They seem not to recognize that much more variation arises than becomes characteristic of the species—most is purged by selection. Moreover, useful variants are initially very rare and require random genetic drift or natural selection to increase their frequency; the combinations of multiple genes underlying most evolved characters cannot be formed or propagated except by selection. These are mathematical and empirically demonstrated facts.

Physical and chemical bases for self-organization exist, of course: they are the mechanisms by which molecules selected for a function perform that function. How could enzymes catalyze reactions if they did not fold (self-organize) according to physicochemical principles? But folding an amino

acid chain does not guarantee a functional enzyme: selection is necessary to shape those relatively few amino acid sequences that function usefully.

The authors' main philosophical argument is that natural selection can be only a correlation rather than a cause, because it cannot distinguish between strongly correlated features or properties. In practice, of course, evolutionary and other biologists separate the effects of correlated traits all the time, sometimes by statistical analysis (few traits are perfectly correlated) and often by manipulative experiments, including many made possible by modern genetic methods such as knockouts. Much of biology consists of determining the functional roles of the correlated elements of complex systems (e.g., biochemical pathways). This reductionist approach differs from Fodor and Piattelli-Palmarini's thoroughgoing

holism, which would render much of biology futile and which is certainly not supported by evolutionary data. Mosaic evolution, the relatively independent evolution of individual traits and genes, may not be universal, but it is very common.

Fodor and Piattelli-Palmarini also argue that the “fit” (adaptation) of organisms to their ecological circumstances is not explained by natural selection because “adaptation” is a tautology: An organism must fit its ecology (or niche). If it were phenotypically other than what it is, it would fit a different ecology. What “explains” the fit is natural history, a multitude of stories such as how the structure of a woodpecker enables it to excavate insects from wood. But natural history “isn't a theory of evolution; it's a bundle of evolutionary scenarios.” Apparently, natural history makes an evolutionary theory of adaptation unnecessary—especially because (they claim) phenotypic evolution is guided not by environmental selection but by endogenous forces. Presumably, then, the authors suppose that the features of, say, woodpeckers (e.g., chisel-like bill, barbed tongue, and elongated hyoid apparatus) arose as a developmental package and that the new phenotype somehow “knew” that it should excavate wood. Such a hopeful-monster scenario is inconsistent with all genetic and other biological knowledge. No known mechanism besides natural selection can account for the assembly of such a combination of genetically complex features. The fit of organism to environment (or “ecol-

The reviewer is at the Department of Ecology and Evolution, Stony Brook University, 650 Life Sciences Building, Stony Brook, NY 11794-5245, USA. E-mail: futuyma@life.bio.sunysb.edu

ogy”), contra Fodor and Piattelli-Palmarini, requires an explanation of its origin.

The diversity of organism-environment relationships reflects the principle that the fitness of phenotypes is highly context-dependent. Thus, the authors argue, there cannot be a universal theory of natural selection, for no general relationship of phenotype to fitness can be specified. But the same might be said of many other research programs. For example, the effect of an enzyme is highly context-dependent, so Fodor and Piattelli-Palmarini presumably would not expect any successful theory in biochemistry. In any case, evolutionary biologists have developed many causal theories, or models, of natural selection on particular classes of traits. We have highly developed theories of the evolution of life history traits, sex ratio, certain principles of form and function, and much more.

These theories of natural selection work: they successfully predict research outcomes.

John Werren predicted and experimentally confirmed that the first of two female parasitic wasps who lay eggs in a host insect lays a more female-biased brood than the second (2). No such prediction could be made without selection theory. Among countless other examples, the pattern of variation in DNA sequences that betokens a “selective sweep” of an advantageous mutation was predicted years before such data could be obtained. Natural selection theory makes successful predictions across a huge range of biological phenomena, and it inspires countless fruitful research programs. What more can one ask of a theory? Contrast that with the ludicrous analogy with which Fodor and Piattelli-Palmarini end: “organisms ‘catch’ their phenotypes from their ecologies in something like the way that they catch their colds from their ecologies.” They helpfully explain that the similarity consists of there being both environmental and endogenous instrumental variables. I look forward to reading about the

research that this formulation will inspire.

Mayr once wrote that “Evolution seems to be a subject on which everybody thinks he is qualified to express an expert opinion” (3). Fodor and Piattelli-Palmarini show little familiarity with the vast literature on genetic variation, experimental analyses of natural selection, or other topics on which they philosophically expound. They are blithely agnostic about the causes of evolution and apparently uninterested in fostering any program of research. Because they are prominent in their own fields, some readers may suppose that they are authorities on evolution who have written a profound and important book. They aren’t, and it isn’t.

References

1. E. Mayr, *Animal Species and Evolution* (Harvard Univ. Press, Cambridge, MA, 1963).
2. J. H. Werren, *Science* **208**, 1157 (1980).
3. E. Mayr, *Toward a New Philosophy of Biology: Observations of an Evolutionist* (Harvard Univ. Press, Cambridge, MA, 1988).

10.1126/science.1189090

NEUROSCIENCE

Meaning-Making Neurons

Michael Shermer

Twice I have spoken at the TED (Technology, Entertainment, Design) conference. Twice I have begrudgingly agreed to the strictly enforced 18-minute talk format—grumbling that “ideas worth spreading” (the TED motto) could not possibly be conveyed in such a constrained format. And twice have I been proven wrong. With discipline and diligence you really can say something of substance in a tight space, and more than 200 million downloads of endlessly entertaining and educational videos (1) prove the principle of pithiness.

In *The Brain and the Meaning of Life*, philosopher, psychologist, and computer scientist Paul Thagard (University of Waterloo) has elegantly employed the pithiness principle. He offers a tightly reasoned, often humorous, and original contribution to the emerging practice of applying science to areas heretofore the province of philosophers, theologians, ethicists, and politicians: What is reality and how can we know it? Are mind and brain one or two? What is the source of the sense of self? What is love? What is the

difference between right and wrong, and how can we know it? What is the most legitimate form of government? What is the meaning of life, and how can we find happiness in it? Thagard employs the latest tools and findings of science in his attempts to answer these (and additional) questions. He briefly reviews how others have addressed them in the past. And he discusses how a scientific worldview can inform one’s analysis and in some cases fully answer the questions—at least to the satisfaction of those of us who take a strictly materialist and naturalist perspective.

Yes, there is a point of view here, and well

The Brain and the Meaning of Life

by Paul Thagard

Princeton University Press, Princeton, NJ, 2010. 292 pp. \$29.95. ISBN 9780691142722.

there should be. When Henry Fawcett commented to Charles Darwin that some scientists found Darwin too theoretical and believed that he should just let the facts speak for themselves, Darwin responded: “How odd it is that anyone should not see that all observation must be for or against some view if it is to be of any service” (2). Thagard’s perspective is that of cognitive neuroscience. He wants to bore into the brain to add a layer of more objective analysis.

Take love, as Thagard does in a concise six pages. He notes that when you gaze upon the face of your lover, the ventral tegmental area and the nucleus accumbens—both rich in dopamine receptors and associated with extremely positive feelings similar to those in cocaine addiction—become quite active. He remarks how the hormone oxytocin increases feelings of attachment between people. These findings and numerous others that he mentions support his model of emotional consciousness, “emocon” (3). That conceptual model sketches how different areas of the brain “interact to produce emotions as the result of both cognitive appraisal and bodily perception.” In it, external stimuli (such as the sight of



The reviewer is at Claremont Graduate University and *Skeptic* magazine, Post Office Box 338, Altadena, CA 91001, USA. E-mail: mshermer@skeptic.com

your loved one) are input through the senses (sight, smells, touch) to the thalamus, which in turn stimulates both brain states and bodily states (increased heart rate and blood pressure, rapid breathing and flushed skin, and so forth). A network of mutual interactions among the amygdala, the insula, and various parts of the prefrontal cortex integrates bodily perceptions and cognitive appraisal. Thus, the base emotions from the amygdala (lusty passion) are linked to the higher cognitive functions of the cerebral cortex (assessment of the relationship).

How does all this get coordinated into a single feeling that we call love? Our dualistic intuitions tell us that there must be a mind that knows what the brain is doing, or some brain module that coordinates all processes into a single self, or some sort of central processing homunculus that sits at a neural switchboard. Not so, says Thagard: "There is no central processor that coordinates all the results and yields a decision. Rather, the brain's reaction to a scary face or other sensory stimulus comes about through the dynamic interaction of external sensory perception, internal sensory perception, cognitive appraisal, and positive and negative valuation." But from where does the sense of a single entity arise? Reciprocal feedback systems: "Note that the connections between brain areas in the ... model are reciprocal, based on neural evidence that there is extensive feedback between neural populations in each pair of regions."

Whether or not reciprocal feedback systems can properly account for such subjective qualia states as love (or for the "self") is highly debatable. Still, Thagard is to be commended for proposing a testable hypothesis and providing evidence in support of it that can be easily accessed by both scientists and general readers. On the subjective feeling of happiness, for example, he cites data gathered by social psychologist Sonja Lyubomirsky on what makes people happy (4). Many things do, among them: expressing gratitude, cultivating optimism, avoiding overthinking and social comparison, practicing acts of kindness, nurturing social relationships, developing strategies for coping, learning to forgive, increasing flow experiences (in which one is absorbed in an activity), savoring life's joys, committing to your goals, practicing religion and spirituality, and taking care of your body through physical activity. Thagard is mildly dismissive of religion and spirituality, but he need not be—just broaden the category to include any activity that generates a sense of awe and transcendence. (For me, that comes from visiting astronomical observatories, fossil quarries, or geological formations, all

BROWSINGS

Caral: La Primera Civilización de América/The First Civilization in the Americas.

Ruth Shady; photographs by Christopher Kleihege. CK Photo (distributed by the University of Chicago Press), Chicago, 2010. 168 pp. \$125, £81. ISBN 9789972337925.

Nearly 5000 years ago, a complex society developed in the desert valleys of the central Peruvian coast. At Caral, comprehensive excavations directed by Shady have uncovered a 65-ha complex of pyramids, smaller platforms, sunken semicircular plazas, public buildings, and residences. This large-format volume presents nearly 200 color photos by Kleihege that depict the site, structures, surrounding landscapes, clay statuettes, and crafted objects (below, stone monolith in front of Pyramid Galería).



of which lead to the contemplation of deep time and the humbling sense of insignificance before the vastness of the cosmos.)

Toward the end of *The Brain and the Meaning of Life*, Thagard dares to employ an objective standard to answer the question "What kind of government should countries have?" Because he gives the topic less than three full pages, political scientists will certainly feel that their field has been short-changed. And while I agree with his conclusion that the current form of government most likely to satisfy human needs is "a liberal democracy operating in a capitalist economic system," I take issue with his subsequent qualification, "with substantial state support for education, health care, and other egalitarian social requirements." Nonetheless, I applaud Thagard's approach of bringing to bear on the question two data sets: the United Nations Human Development Index (5) (which rates 177 countries on how well they provide their citizens with "a long and healthy life, education, and a decent standard of living") and yearly surveys, since 1981, of subjective well-being (happiness) (6). Iceland,

Canada, Ireland, the Netherlands, and Switzerland appeared near the top in both lists. It is true that two data sets do not a political science make, but Thagard's concluding remarks in this section are a model of scientific caution and skepticism: "We should also not rule out the possibility that some form of government not currently practiced might actually be better for meeting vital human needs than those now in operation. Perhaps future social experiments will find creative new ways of governing states that will be more effective than those now observed." Although we cannot implement such experiments in the name of science, if they do happen, scientists should be the first in to record the results.

References

1. www.ted.com.
2. Letter, C. R. Darwin to H. Fawcett, 18 September 1861; www.darwinproject.ac.uk/entry-3257.
3. P. Thagard, B. Aubie, *Conscious. Cogn.* **17**, 811 (2008).
4. S. Lyubomirsky, *The How of Happiness: A Scientific Approach to Getting the Life You Want* (Penguin, New York, 2008).
5. <http://hdr.undp.org/en/statistics/>.
6. www.worldvaluessurvey.org/happinessrends.

10.1126/science.1189752

SCIENCE AND SOCIETY

Testing Time for Climate Science

Sheila Jasanoff

On 31 March 2010, a British parliamentary committee exonerated Philip D. Jones, director of the Climatic Research Unit (CRU) at the University of East Anglia, of personal wrongdoing in his conduct and management of research. Climate science fared less well. The Science and Technology Committee concluded in its report that the focus on a single individual had been misplaced: “we consider that Professor Jones’s actions were in line with common practice in the climate science community” (1). Those practices included routine refusals to share raw data and computer codes. The committee judged that this had to change and that all future raw data and methodological work should be publicly disclosed.

In early 2009, few would have predicted that climate science was headed for a public trial or public embarrassment. The Intergovernmental Panel on Climate Change (IPCC), the world’s chief provider of scientific knowledge about the climate, enjoyed a pristine reputation. With nearly two decades of work and four assessment reports to its name, the IPCC seemed to have quelled the doubts of many skeptics. A growing scientific consensus accepted the anthropogenic causes of climate change (2). Added validation came when the IPCC shared the 2007 Nobel Peace Prize with former Vice President Al Gore. President Barack Obama earned worldwide commendation when he signaled that America was at last willing to act on the IPCC’s painstakingly assembled knowledge.

The ground shifted dramatically in November 2009 with the event that became known as “climategate” (3). A hacker entered the CRU’s computer system and disclosed some 1000 private e-mails and 3000 documents. Some showed climate scientists apparently fudging data to exaggerate the effects of warming. Words like “trick” and “hide,” referring to modelers’ techniques of representing data, were seized upon as signs that CRU was purposefully distorting results to support its claims. Other messages suggested that scientists were reluctant to make raw data available to known critics and had tried to keep unfriendly papers from publication in peer-reviewed journals. In the ensu-

ing uproar, the credibility of climate science suffered. A poll conducted in February 2010, found a 30% drop over 1 year in the percentage of British adults who believe climate change is “definitely” real (4).

In a time when global policy increasingly depends on scientific knowledge, the CRU’s plight is not good news for science or society. What can be done to guard against such setbacks and to rebuild public faith in the credibility of climate science? A half-century of scientific advising holds some lessons.

From Integrity to Accountability

Scientific progress has always depended on credibility and trust. To build new knowledge, scientists have to be able to take each other’s findings at face value. If every claim needed to be verified before others could act on it, research would grind to a halt. English experimental scientists in the 17th century set out to perfect, not only their methods of inquiry, but also the techniques of communication that would enhance credibility. For example, the adoption of an impersonal writing style increased the appearance of objec-

Climate science needs better ways of accounting for itself to the jury of the world.

tivity (5). As in the law, fact-finding in science also called on witnesses to validate new claims. The sociologist R. K. Merton attributed the rise of peer review, a form of “organized skepticism,” to scientists’ need for results that could be trusted (6).

In earlier times, it was enough to build trust within a researcher’s community of scientific peers. Disciplines were small and methodologically coherent. Research neither drew heavily on public funds nor profoundly affected public decisions. Today, the circle of stakeholders in science has grown incomparably larger. Much public money is invested in science and, as science becomes more enmeshed with policy, significant economic and social consequences hang on getting the science right. Correspondingly, interest in the validity of scientific claims has expanded to substantially wider audiences. It is not only the technical integrity of science that matters today but also its public accountability.

In the United States, an elaborate legal framework for holding policy-relevant science accountable has been in the making since just after World War II. The 1946 Administrative Procedure Act (Public Law 79-404) required federal agencies to consult with the public before enacting new regulations; at minimum, providing notice and an opportunity to comment. A later milestone was the 1969 National Environmental Policy Act (NEPA) (Public Law 91-190), which called for extensive public inputs. Scoping exercises and hearings designed to solicit information from the public and to explain agency findings became recognized elements of the NEPA process. Many environmental and consumer protection laws now mandate public involvement beyond the requirements of notice and comment. Moreover, administrative decisions can be overturned if an agency does not have adequate scientific and technical evidence or has failed to act reasonably on the basis of available knowledge (7). Under the Federal Advisory Committee Act (FACA) (Public Law 92-463), scientific advisory committees must be fairly balanced and, in the absence of special circumstances, committee meetings and records are presumed to be open to the public.

The rising importance of public accountability is also reflected in growing concern with ethics in science and the proliferation of ethics oversight bodies. Once limited largely



Climate coalition protest rally, Brussels, 2009.

John F. Kennedy School of Government, Harvard University, Cambridge, MA 02138, USA. E-mail: sheila_jasanoff@harvard.edu

to concern for the welfare of human and animal subjects, today, ethics covers a wide array of issues across many emerging areas of science and technology, including stem cell research, nanotechnology, computer science, and the neurosciences. It is no longer enough to establish what counts as good science; it is equally important to address what science is good for and whom it benefits.

A 1983 and a 1996 report of the National Research Council bookended the turn from integrity to accountability. The first (8) recommended that the largely scientific exercise of risk assessment should be separated as far as possible from the political and value-laden task of risk management. The chief purpose was to protect science against possible biases. The second (9) concluded that risk analysis should be seen as an intertwined analytic-deliberative process, requiring repeated public consultation even in the production and assessment of scientific knowledge. Here, there was recognition that public consultation improves the quality and acceptability of expert judgments.

Science today has to meet a series of public expectations, not only about its products but also about its processes and purposes. The credibility of climate science has to be evaluated in this context of heightened demand for accountability. Accountability can be seen as a three-body problem, with each interacting component posing special problems for climate science.

A Three-Body Problem

The individual scientist or expert. In any professional activity where truth-telling counts—whether in law, accounting, engineering, medicine, or science—practitioners must be held to high standards of honesty and integrity. In science, peer review partly serves this purpose, weeding out dishonesty and misrepresentation along with mistaken or inconclusive results. Of course, the scientific community has experienced many episodes of misconduct (10), but there is often broad agreement on what constitutes deviant behavior, and publics by and large have reason to trust science's self-correcting practices.

Scientific knowledge. This body is organized into disciplines or into well-defined, topically focused areas of inquiry. Reliable bodies of knowledge are built on theories and methods that have wide currency among practitioners. Again, peer review serves a crucial legitimating function by maintaining rigor, coherence, and integrity in the development of a field's research frontiers. Peer review also demarcates work that is considered acceptable from work that is not (11). In many areas

of science, the ongoing work of peer criticism is enough to ensure a field's credibility to the outside world.

Committees that translate scientific findings into policy-relevant forms. This third body is increasingly important in modern democracies and frequently combines knowledge and skills from experts in different fields and contexts—for example, science and engineering, universities and industry, and bench and clinic. Their authority derives in part from individual members' impartiality and sound judgment and in part from the views they collectively represent, as required in the United States by FACA. Scientific advisory committees have dealt with the demand for accountability far longer than scientists who never did the work of translating science for policy. In most Western countries, expert advisers are required to explain their judgments to audiences outside, as well as within, their own research communities (12).

Implications for Climate Science

Standards of individual good behavior are especially difficult to identify and enforce in evolving scientific domains with underdeveloped histories of accounting to external audiences. Divergent national traditions of openness and confidentiality present additional hurdles for climate scientists (13), who are involved in international, as well as interdisciplinary, consensus-building. As the UK inquiry on the hacked CRU e-mails revealed, some data relied on by climate scientists had been obtained from national governments under nondisclosure agreements. The parliamentary committee conducted, in effect, a process of post hoc standard-setting when it concluded that the climate science community should have followed more open practices of publication and disclosure.

The sciences represented by IPCC Working Group I do not share common principles for such basic tasks as visualizing data, interpreting anomalies, representing uncertainty, data-sharing, or public disclosure. That such disparate communities have come to agree on the causes, size, and scope of the climate problem, through iterative rounds of assessment, may be taken as strong evidence of reliability. At the same time, the very fact that judgment has been integrated across many fields leaves climate science vulnerable to charges of groupthink and inappropriate concealment of uncertainties.

Though intergovernmental in name, the IPCC is subject to none of the legal or political requirements that constrain, but also legitimate, national expert committees. The IPCC has invented its own procedures, includ-

ing extensive and sophisticated peer review. These methods are good enough to satisfy many scientists, but they rest on traditions of scientific, rather than public, accountability. Yet the IPCC performs a mix of functions—part scientific assessment, part policy advice, and part diplomacy—that demand external, as well as internal, accountability.

These problems suggest that it will not be enough for climate scientists to be still more scrupulous and transparent toward their peers. Adding more new forms of expertise may increase the credibility of the field (14), but it will not fully address the third component of accountability, which involves relations between science and its publics.

Creating accountability practices that work at a supranational level will be neither straightforward nor easy. Administrative procedures mostly operate within nation states, and there is no higher court where science can account for itself to the world. However, the IPCC has demonstrated that it can learn and change in its methods of representing science to scientists. That ingenuity should now be directed toward building relationships of trust and respect with the global citizens whose future climate science has undertaken to predict and reshape.

References and Notes

1. Science and Technology Committee, U.K. Parliament, "The disclosure of climate data from the Climatic Research Unit at the University of East Anglia," Eighth Report of Session 2009-10 (HC 387-I, Stationery Office, London, 2010), p. 3.
2. N. Oreskes, *Science* **306**, 1686 (2004).
3. A. C. Revkin, *New York Times*, 21 November 2009, p. A1.
4. J. Jowit, *Guardian*, 23 February 2010, p. 9.
5. S. Shapin, S. Schaffer, *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life* (Princeton, NJ: Princeton Univ. Press, 1985).
6. R. K. Merton, *The Sociology of Science: Theoretical and Empirical Investigations* (Univ. of Chicago Press, Chicago, 1973), chap. 13.
7. For example, *Massachusetts v. EPA*, 549 U.S. 497 (2007), held that EPA had to give more persuasive reasons for its decision not to regulate greenhouse gases as air pollutants under the Clean Air Act.
8. National Research Council, *Risk Assessment in the Federal Government: Managing the Process* (National Academy Press, Washington, DC, 1983).
9. P. C. Stern, H. V. Fineberg, Eds., *Understanding Risk: Informing Decisions in a Democratic Society* (National Academies Press, Washington, DC, 1996).
10. W. Broad, N. Wade, *Betrayers of the Truth: Fraud and Deceit in the Halls of Science* (Simon & Schuster, New York, 1983).
11. T. F. Gieryn, *Cultural Boundaries of Science: Credibility on the Line* (Univ. of Chicago Press, Chicago, 1999).
12. S. Jasanoff, *The Fifth Branch: Science Advisers as Policymakers* (Harvard Univ. Press, Cambridge, MA, 1990).
13. R. Brickman, S. Jasanoff, T. Ilgen, *Controlling Chemicals: The Politics of Regulation in Europe and the United States* (Cornell Univ. Press, Ithaca, NY, 1985).
14. A. C. Revkin, "East Anglia's climate lessons," *Dot Earth blog*, *New York Times*, 14 April 2010.

10.1126/science.1189420

IMMUNOLOGY

Tumor Immune Evasion

Carlene L. Zindl¹ and David D. Chaplin²

Many types of human tumors can suppress the immune system to enhance their survival. Some tumor cells escape immune detection by decreasing the expression of certain antigen-presenting proteins at their surface, rendering them invisible to cytotoxic T lymphocytes (1). But more often, tumors secrete proteins that inhibit effector T cell responses and promote the production of regulatory T cells that suppress immune responses (2). On page 749 of this issue, Shields *et al.* (3) identify another mechanism by which tumors deceive the immune system. Certain melanomas can reorganize their stromal microenvironment (the supportive connective tissue) into structures similar to lymphoid tissue of the immune system. This ingenious reconstruction recruits and maintains immune regulatory cells that promote tolerance and tumor progression.

Lymphoid tissue inducer (LTi) cells contribute to the organogenesis of secondary lymphoid tissues (lymph nodes and Peyer's patches) during mammalian development (4, 5). These cells express members of the tumor necrosis factor (TNF) cytokine family—TNF, lymphotoxin (LT $\alpha_1\beta_2$), and TNF-related activation-induced cytokine

(TRANCE). Together, these factors boost the expression of chemokines and adhesion molecules by stromal “organizer” cells during fetal development (6). As the stromal cells (tissue supporting cells) of the lymphoid organ develop, they express adhesion molecules and chemokines (such as CXCL13 and CCL21) that support their stable interactions with fetal LTi cells, as well as recruit LTi cells, B cells, and T cells into the developing tissue (7). Ultimately, this localized stromal cell activation results in an organized, compartmentalized lymphoid organ that regulates immune responses.

In adult mammals, LTi cells direct the reorganization of stromal cells into tertiary lymphoid structures (lymphoid tissue–like structures found at nonlymphoid sites) under conditions of persistent inflammation or infection (8, 9). Shields *et al.* show that mouse and human melanomas expressing the chemokine CCL21 can recruit LTi cells. This results in reorganization of the tumor's stroma and the recruitment of CD4⁺ regulatory T cells, myeloid-derived suppressor cells, and other leukocytes (see the figure). Because other signals in addition to CCL21 contribute to secondary and tertiary lymphoid tissue development, it may be that other tumor types that use these signaling pathways may induce stroma reconstruction as well.

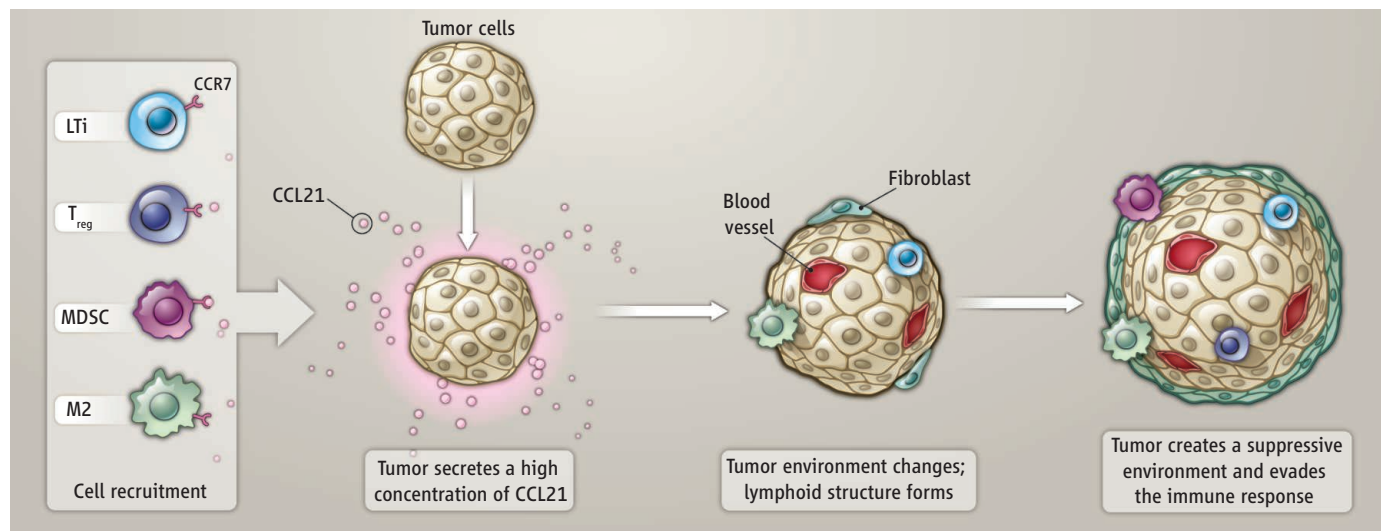
Secondary and tertiary lymphoid tissues provide a favorable environment for activating

Malignant cells can induce the formation of lymphoid tissue–like structures that help the tumor evade host immunity.

humoral and cellular immunity. But if a tumor thrives when it escapes immune responses, then why would it create surroundings with characteristics of an immune system tissue? The association of lymphoid structures with tumors underscores the fact that lymphoid tissues can both activate and down-regulate immune effector pathways. For example, stromal cells in lymph nodes present antigen to CD8⁺ T cells in a way that induces tolerance rather than activation (10). Additionally, lymphoid tissues provide an environment in which naïve T cells, in the presence of transforming growth factor- β (TGF- β), become regulatory T cells, further favoring the suppression of effector T cell functions. Thus, by mimicking the functions of secondary or tertiary lymphoid tissues, lymphoid tissue–like stromal structures created by tumors can promote immune tolerance and suppression.

Do melanoma cells directly block or reduce the differentiation of naïve T cells into effector T cells, or do they act indirectly by recruiting immunoregulatory antigen-presenting cells into the lymphoid tissue–like structures? Shields *et al.* observed that B cells were not recruited to the site of melanoma cell accumulation. Because B cells can prime T cells in secondary lymphoid tissues, perhaps their absence reduces immune activation within the tumor-induced stroma. Alternatively, TGF- β produced in the tumor may alter local macrophage populations, resulting

¹Department of Pathology, University of Alabama at Birmingham, Birmingham, AL 35294, USA. ²Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294, USA. E-mail: dchaplin@uab.edu



Lymphoid stroma formation. Mouse tumor cells expressing the chemokine CCL21 recruit lymphoid tissue inducer (LTi) cells, regulatory T (T_{reg}) cells, and myeloid-derived suppressor cells (MDSCs). LTi cells direct the reorganization

of local stroma into lymphoid tissue–like structures that support T_{reg} cells and MDSCs and lead to immune tolerance of the tumor. M2 macrophages secrete extracellular matrix proteins that promote stromal stability.

CREDIT: Y. GREENMAN/SCIENCE

in a switch from classically activated, phagocytic, and proinflammatory M1 macrophages to alternatively activated, poorly phagocytic, and anti-inflammatory M2 macrophages that secrete extracellular matrix proteins (11, 12). This raises the question of whether factors normally produced by tumors alter the type of antigen-presenting cells in the tumor microenvironment in ways that influence the immune response and promote stromal stability. Understanding how tumor-induced changes in the stroma influence different immune cells will be important for determining ways to modulate these structures.

The involvement of LT α i cells in the lymphoid tissue-like structures observed by

Shields *et al.* suggests that the cytokine LT α β ₂ may act to increase the expression of the chemokine CCL21. This may be important because LT α β ₂-dependent structures are generally plastic. Thus, the lymphoid tissue-like structures should be reprogrammable by modulating LT α β ₂ signaling, as has been observed in stromal compartments of the mouse spleen (13). Understanding the extent to which tumor-induced lymphoid tissue-like structures are plastic will go a long way toward determining if blocking CCL21 and/or LT α i cell function can disrupt immune tolerance of the tumor-induced structures, thereby releasing host immunity to aid in eliminating the malignant cells.

References

1. M. Meissner *et al.*, *Clin. Cancer Res.* **11**, 2552 (2005).
2. E. M. Shevach, *Nat. Med.* **10**, 900 (2004).
3. J. D. Shields, I. C. Kouritis, A. A. Tomei, J. M. Roberts, M. A. Swartz, *Science* **328**, 749 (2010); published online 25 March 2010 (10.1126/science.1185837).
4. R. E. Mebius, P. Rennert, I. L. Weissman, *Immunity* **7**, 493 (1997).
5. H. Yoshida *et al.*, *Int. Immunol.* **11**, 643 (1999).
6. D. Kim *et al.*, *J. Exp. Med.* **192**, 1467 (2000).
7. K. M. Ansel *et al.*, *Nature* **406**, 309 (2000).
8. F. Aloisi, R. Pujol-Borrell, *Nat. Rev. Immunol.* **6**, 205 (2006).
9. T. Cupedo *et al.*, *Immunity* **21**, 655 (2004).
10. J. W. Lee *et al.*, *Nat. Immunol.* **8**, 181 (2007).
11. A. Mantovani *et al.*, *Trends Immunol.* **23**, 549 (2002).
12. A. Sica *et al.*, *Cancer Lett.* **267**, 204 (2008).
13. C. L. Zindl *et al.*, *Immunity* **30**, 408 (2009).

10.1126/science.1190310

PLANETARY SCIENCE

Cometary Dust in the Laboratory

Larry R. Nittler

Much of our understanding of early solar system history comes from the laboratory study of extraterrestrial materials, especially meteorites from the asteroid belt, as these are by far the most readily available. However, it has long been recognized that icy bodies in the outer solar system might contain a better-preserved record of the earliest stages of solar system formation than can be found in even the most primitive asteroidal meteorites. This is supported by laboratory studies of tiny (<50 μ m diameter) meteorites—interplanetary dust particles or IDPs—collected by aircraft in the stratosphere. Thought to originate at least in part from comets, IDPs exhibit highly primitive features (1, 2). Two reports, one by Matzel *et al.* (3) and one by Duprat *et al.* (4) on page 742 of this issue, take advantage of the availability of relatively less studied primitive materials—samples obtained by the Stardust mission from comet Wild 2 and Antarctic micrometeorites (AMMs)—to glean information about the early solar system.

In 2006, NASA's Stardust returned solid samples similar in size to IDPs (5) (see the

Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, NW, Washington, DC 20015, USA. E-mail: lnittler@ciw.edu



figure, panel A). Much less attention has been paid to intermediate-sized AMMs obtained by filtering melted Antarctic snow and/or ice (6). AMMs are in the size range (hundreds of micrometers) that dominates the flux of extraterrestrial materials incident on Earth, but particles of this size are the most susceptible to strong modification or destruction by heating during atmospheric entry. However, prior work has shown that some AMMs

Comparison of micrometeorites collected in space and the Antarctic snow provides insight into the early solar system.

Sample collection. (A) Dust sample captured in an aerogel matrix during NASA's Stardust mission to meet with comet Wild 2; the "Coki" grain studied by Matzel *et al.* (3) is circled. (B) Researchers collecting snow near CONCORDIA research station at Dome C, Antarctica (75°S, 123°E). Melting and sieving of this snow provided numerous micrometeorites (e.g., inset), including the ultracarbonaceous ones discussed by Duprat *et al.* (4).

have survived their fall to Earth relatively unscathed (6) and sample a broader range of parent bodies, likely including comets (7), than do the larger meteorites.

A major surprise came from the initial analysis of the Stardust samples from Wild 2. The cometary dust was similar in many ways to materials found in primitive meteorites (8), in particular the abundance of high-temperature phases, which must have formed in the inner regions of the solar system and been subsequently transported outside the orbit of Neptune. The high-temperature material included grains similar to (but much smaller than) calcium- and aluminum-rich inclusions (CAIs) (9), the oldest known solids in meteorites, and chondrules (10), round silicate objects that dominate many primitive meteorites. Matzel *et al.* report the first search for extinct ^{26}Al in a grain from Wild 2, nicknamed Coki. The radioisotope ^{26}Al (half-life = 730,000 years) was homogeneously distributed in the early inner solar system (11), and its abundance in various meteoritic samples, inferred from its decay product (stable ^{26}Mg), can thus be used to infer relative chronological information with high time resolution. Samples formed

CREDIT: (PANEL A) DAVID JOSHIWAK/UNIV. WASHINGTON; (PANEL B) JEAN DUPRAT/CNRS

earlier have higher inferred $^{26}\text{Al}/^{27}\text{Al}$ ratios.

Coki is 5 μm in diameter and is primarily composed of the feldspar mineral anorthite, whose Al/Mg ratio is high enough for ^{26}Al to be detected if present at the same level as seen in many meteoritic CAIs, despite its small size. However, the authors found no evidence for extinct ^{26}Al in this grain, and their upper limit on the initial $^{26}\text{Al}/^{27}\text{Al}$ ratio indicates that it must have formed at least 1.7 million years after the oldest CAIs. Thus, large-scale transport of material from the inner to the outer solar nebula must have taken place over a time scale of millions of years, and Wild 2 itself must have accreted after this time. This observation supports the conclusion that comets are not simply collections of interstellar or early-formed primitive materials.

A very different type of material was studied by Duprat *et al.* The AMMs in their study were collected by melting and filtering snow that fell in the mid-20th century near the French-Italian CONCORDIA station (see the figure, panel B). Among the recovered particles were some fluffy, fine-grained ones with much higher carbon contents than seen in other primitive materials, including meteorites, most IDPs, and Wild 2 samples. Detailed study of these “ultracarbonaceous” AMMs (UCAMMs) revealed that the carbon is in the form of a poorly ordered organic material, similar to that seen in the most

primitive meteorites and IDPs. Most interesting is the discovery that the organic matter is highly enriched in deuterium, with D/H ratios up to 30 times terrestrial values. Such high D/H ratios have been seen in carbonaceous chondrite meteorites (12) and IDPs (13), but typically only as submicrometer-to micrometer-sized grains (“hot spots”). In contrast, organics in the UCAMMs show extreme D/H ratios extending over tens of square micrometers. The origin of D enrichments in primitive organic matter is a matter of controversy, with interstellar, protostellar, and parent-body processes all possibly playing a role.

Embedded in the D-rich organic matter were crystalline and amorphous minerals similar to those found in IDPs and Wild 2 samples. The primitive nature of the UCAMMs, especially the high abundance of D-rich organic matter, suggests a connection to icy bodies in the outer solar system, with comets being the most likely to supply dust to Earth-crossing orbits. The 1986 Giotto mission to comet Halley identified large numbers of carbon-rich particles intermixed with silicate minerals [so-called CHON particles (14)], and the UCAMMs may well be similar objects. However, similar extremely D-rich carbonaceous materials have not yet been identified in the only unambiguously cometary samples, those from Wild 2 (15). This may be due in part to

sampling biases, but likely also reflects substantial chemical diversity among comets.

Detailed laboratory comparisons of materials from diverse parent bodies from throughout the protosolar disk can provide insights into processes including the timing of grain formation and transport, as well as the origin and distribution of organic matter. Moreover, the very high carbon contents of UCAMMs may well have profound implications for the original delivery of organic molecules to the early Earth, with possible consequences for the earliest prebiotic chemistry.

References

1. J. P. Bradley, *Science* **265**, 925 (1994).
2. H. Busemann *et al.*, *Earth Planet. Sci. Lett.* **288**, 44 (2009).
3. J. E. P. Matzel *et al.*, *Science* **328**, 483 (2010); published online 25 February 2010 (10.1126/science.1184741).
4. J. Duprat *et al.*, *Science* **328**, 742 (2010).
5. D. Brownlee *et al.*, *Science* **314**, 1711 (2006).
6. M. J. Genge, M. M. Grady, R. Hutchison, *Geochim. Cosmochim. Acta* **61**, 5149 (1997).
7. D. Nesvorný *et al.*, *Astrophys. J.* **713**, 816 (2010).
8. H. A. Ishii *et al.*, *Science* **319**, 447 (2008).
9. M. E. Zolensky *et al.*, *Science* **314**, 1735 (2006).
10. T. Nakamura *et al.*, *Science* **321**, 1664 (2008).
11. J. Villeneuve, M. Chaussidon, G. Libourel, *Science* **325**, 985 (2009).
12. H. Busemann *et al.*, *Science* **312**, 727 (2006).
13. S. Messenger, *Nature* **404**, 968 (2000).
14. M. E. Lawler, D. E. Brownlee, *Nature* **359**, 810 (1992).
15. K. D. McKeegan *et al.*, *Science* **314**, 1724 (2006).

10.1126/science.1187725

ENGINEERING

The Impact of Tellurium Supply on Cadmium Telluride Photovoltaics

Ken Zweibel

For decades, the material associated with photovoltaic (PV) cells has been silicon. However, after many years of development, cadmium telluride (CdTe) PV modules have become the lowest-cost producer of solar electricity, despite working at lower efficiency than crystalline silicon cells. CdTe sales are growing rapidly, but there is concern about projecting hundredfold increases in power production relative to current production with CdTe PV modules. One reason is that Te, a humble nonmetal that is actually abundant in the universe, is as rare as many of the precious metals recovered from Earth’s crust (1). Fur-

thermore, current technology now uses Te at rates that are substantial fractions of its supply. Here, I argue that the long-term potential for CdTe PV modules need not be bleak, given realistic developments in module technology and Te recovery.

That Te supply is even an issue in thinking about the future of renewable energy results from recent decreases in production costs of the PV module—the deployable device, or large-area aggregation of solar cells, that contains the active PV materials and delivers current. CdTe module production costs have dropped from over \$2/W in 2004 to \$0.84/W in 2010, the lowest in photovoltaics (2). A key advantage of CdTe for thin-film devices is that it can be deposited rapidly. For other thin-

Better optical designs and enhanced recovery of tellurium may boost the potential for large-scale energy production from thin-film cadmium telluride solar cells.

film PV materials, the vapor-phase composition must be carefully adjusted and controlled, which slows the process and adds to costs. CdTe can be deposited at rates of micrometers per minute over large substrates, versus nanometers per minute for amorphous silicon.

These cost reductions bring PV closer to competitiveness with current power generation cost. In the United States, the approximate cost would be as low as 15 cents per kilowatt-hour (c/kWh) for parts of the Southwest that have the most available sunlight. For the rest of the USA, it would be 20 c/kWh (retail prices for electricity in the USA are now about 10 c/kWh) (3, 4).

Any favorable projections for CdTe PV will be moot if its contribution to power pro-

The George Washington University, Washington, DC 20052, USA. E-mail: zweibel@gwu.edu

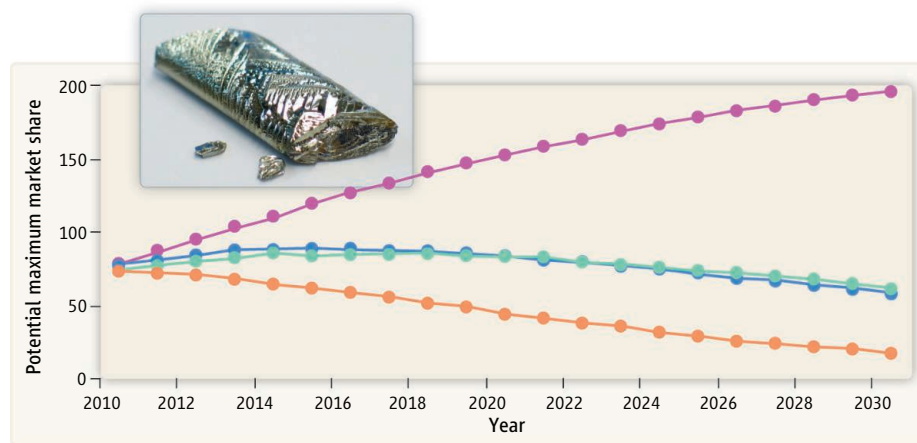
duction is severely limited by Te supply. For the present technology, generating 1 gigawatt (GW) of power requires 91,000 kg (91 metric tons, MT) of Te (a cost of about \$20 million). If all of the PV delivered in 2009–7 GW—had been produced with CdTe, about 640 MT of Te would have been required, which is comparable to its present annual production (5, 6).

If PV is to supply 10% of the projected demand of electricity worldwide in 2030, the per annum growth rate must be 18.5%; for 25% of world electricity, it must be 25%. In 2030, the annual production would require 200 GW/year (at 10% electricity supply) or 670 GW/year (at 25% supply). For the current CdTe modules, 19,000 or 61250 MT of Te per year, respectively, would be needed, equivalent to an increase in supply by a factor of about 40 to more than 100, respectively.

However, future developments may make this challenge less daunting. First, there will likely be some gains made in supply given the inducement of higher Te prices. The crustal abundance of Te is similar (7) to that of platinum, but the actual recovery of Te is more than twice that of platinum. The existing supply is obtained almost exclusively through Te recovery as a by-product of refining copper. Ojebuoboh (8) has described how the supply could expand from 500 to 1500 MT/year by electrorefining more copper to recover more Te. Availability should also increase simply through the historical increase in copper extraction, typically 1 to 3% per year. Also, several other primary metals besides copper have Te impurities. The highest Te concentrations occur in gold deposits, and Te could be refined from zinc and lead (9).

Further, there are already sources of Te as a primary ore. Mines in Mexico (10), China (11), and Sweden (12) have rich bismuth telluride ores, with Te concentrations of almost 20% (9). The challenge with primary ores is that few have been identified, and that the Te must be rich enough to refine alone. This is perhaps about 0.25% concentration, if no by-product minerals can be processed and sold. Finally, there appear to be undersea deposits (7); Te-rich layers (mean concentration of 50 parts per million) coat undersea ferromanganese crusts.

Most importantly, it may also be possible to use much less Te and still maintain module performance. One possibility is to decrease the thickness of the active layer. It does not take 3 μm of CdTe to absorb the solar spectrum. In fact, over much of its absorption range, the CdTe light absorption coefficient is above $10^5/\text{cm}$, that is, it absorbs 63% of available photons in 10^{-5} cm, or 0.1 μm (a



Projecting paths for CdTe photovoltaics. The concerns about Te availability limiting CdTe PV module production assume that the layer thickness will be maintained at 3 μm . Projections of maximum market share attainable are shown based on modest increases in Te production (from 1% growth per year in copper production, its main supply route) and module efficiency (15%), but substantial decreases in CdTe active-layer thickness. The blue and purple lines are market-share maxima if 10% of the world electricity is made from PV in 2030, with layer thicknesses of 0.67 and 0.2 μm , respectively. Similar projections (orange and green) are for 25% world electricity production by PV in 2030. In all cases, with thin enough CdTe, nearly 100% market share might be attained. The inset shows tellurium in its native form.

thickness decrease of a factor of 30).

However, it is not so simple to make such ultrathin CdTe films work well. Two major loss mechanisms for absorbed photons have to be addressed to use CdTe at 0.1- μm thickness. First, 37% of the spectrum would be lost, and the equivalent loss of efficiency would be unacceptable. Second, absorption rises to 10^{-5} cm for higher-energy (shorter-wavelength) light, but longer-wavelength light just above the 1.5-eV band gap of CdTe (the minimum light energy at which it generates a photocurrent) requires thicker CdTe in order to be absorbed. Because 1.5 eV is the

the 0.1- to 0.3- μm range. The mirror material is a combination of a transparent conductive oxide (ZnO) and a silver or aluminum film on top of the back contact.

The rapid deposition of thicker CdTe layers help to minimize manufacturing costs, which presently are greater than costs associated with Te. When Te becomes the limiting cost factor, a shift to thinner CdTe will likely require slower deposition rates to gain greater control of layer thickness.

The other major lever in reducing the amount of Te needed is to increase efficiency. CdTe modules are 11% efficient, and the best cells are 16.5%. Theoretical efficiencies for a single junction cell top out at 33%, and the band gap for CdTe is very close to the optimum for taking advantage of the solar spectrum. Conservative long-term goals for the efficiency of CdTe modules tend to be about 15%. The table shows how the amount of Te required per gigawatt of energy generation would change for different efficiencies and layer thicknesses.

What does this mean in terms of the long-term contribution of CdTe to PV energy production? The growth of PV deployment will take decades. During that time, Te refining, Cu extraction, CdTe module efficiency, and CdTe layer thickness can be improved. The projections shown in the figure have not counted on more Te from new bismuth telluride ores, undersea ridges, or greater refining of non-Cu ores (13). The limit on market share for 10 and 25% production of the world's electricity by PV shown in the figure are for a CdTe module efficiency of 15%. For 10% PV electricity production in 2030, the

POTENTIAL FOR USING LESS Te (IN METRIC TONS) PER GIGAWATT OF PHOTOVOLTAIC OUTPUT			
Efficiency	CdTe layer film thickness		
	3 μm (now)	2/3 μm	0.2 μm
10%	100 MT/GW	22 MT/GW	6.6 MT/GW
15%	67 MT/GW	15 MT/GW	4.4 MT/GW

middle of the solar visible range, too much light would be wasted.

Two commercial PV technologies already deal with these problems—amorphous and thin-film microcrystalline silicon. They use two strategies to trap light inside the cell. One is to deposit a mirror on the back of the cell, just above the metal back contact. The other is to texture the top, transparent conductive oxide contact so that light is bent as it enters the cell (making the first pass longer, and leading to total internal reflection if combined with a back-side mirror). Amorphous silicon, with about the same absorption coefficient as CdTe, is already made with these ancillary components, and layers are in

numbers are encouraging for 0.67- μm layer thickness, and of course better for 0.2 μm . In this case, all modules could be CdTe. For the 25% PV electricity in 2030, the goal could be reached with 0.2- μm layers.

So this analysis brings us back to the question—will CdTe contribute substantially to renewable energy production? Supply and demand will almost certainly remain a key issue in CdTe PV production. However, projections that underestimate CdTe PV at this early stage of development assume a situation for Te supply and use that will be nearly static during the next 20 years. In an area driven by innovation, this grim scenario seems unlikely.

References and Notes

1. B. L. Cohen, *Geochim. Cosmochim. Acta* **48**, 203 (1984).
2. First Solar accounts for almost all production of CdTe modules at this time. Abundant Solar has had some production in 2010. PrimeStar Solar—GE, Q-Cells Calyxo, and Arendi are CdTe PV startup companies that remain largely precommercial.
3. U.S. Energy Information Administration, *Electric Power Monthly*, March 2010, <http://tonto.eia.doe.gov/ftproot/electricity/epm/02261003.pdf>.
4. First Solar has a technical roadmap that would reduce near-term costs by one-third. However, silicon module costs have plunged recently, so that the cost difference between First Solar products and the best of the silicon products, many from China, is now small—perhaps 20% at the system level.
5. M. A. Green, *Prog. Photovoltaics* **14**, 743 (2006).
6. V. Fthenakis, *Renewable Sustainable Energy Rev.* **13**, 2746 (2009).
7. J. R. Hein, A. Koshchinsky, A. N. Halliday, *Geochim. Cosmochim. Acta* **67**, 1117 (2003).
8. F. Ojebuoboh, *World Metallurgy—ERZMETALL*—Heft 1/2008 **61**, 1 (2008).
9. W. C. Cooper, Ed., *Tellurium* (Van Nostrand, New York, 1971), chap. 1.
10. "Mexivada stakes new gold-tellurium property in Mexico," Press Release, 12 June 2008 (www.mexivada.com).
11. Sichuan Apollo Solar T&D Inc., 2008, www.scasolar.com/english/web.asp?id=186 (accessed 11 June 2008).
12. Gold Ore Resources, <http://metalsplace.com/news/articles/24830/gold-ore-announces-positive-tellurium-metallurgical-studies> (2009).
13. It is a mystery why there is so much tellurium in the universe (1) and so little found on Earth. Do the undersea ridges explain the crustal absence? Or is there much more, elsewhere, waiting to be found?
14. The author was a cofounder of PrimeStar Solar. George Washington University Solar Institute is a member of Thin Film PV Partnership and receives financial support from First Solar.

10.1126/science.1189690

NEUROSCIENCE

Epigenetics and Cognitive Aging

J. David Sweatt

Cognitive decline, especially in memory capacity, is a normal part of aging (1). Indeed, the painful reality is that aging-related cognitive decline likely begins when one is in their late 40s. This deterioration is particularly pronounced in declarative memory—the ability to recall facts and experiences—and has been associated with aberrant changes in gene expression in the brain's hippocampus and frontal lobe. However, the molecular mechanisms underlying these changes in gene regulation are not currently known (2, 3). On page 753 of this issue, Peleg *et al.* (4) bolster an emerging hypothesis that changes in the epigenetic modification of chromatin in the adult central nervous system drive cognitive decline.

Chromatin remodeling in the hippocampus is necessary for stabilizing long-term memories (5–8). The relevant molecular mechanisms include DNA methylation and the modification of histone proteins by acetylation, phosphorylation, and methylation. These epigenetic changes involve covalent chemical modifications by enzymes such as histone acetyltransferases and histone deacetylases. Whether alterations in these mechanisms contribute to age-related changes in gene transcription and memory decline are unknown (8).

Peleg *et al.* found that aged mice exhibit a disruption of memory-associated activity-

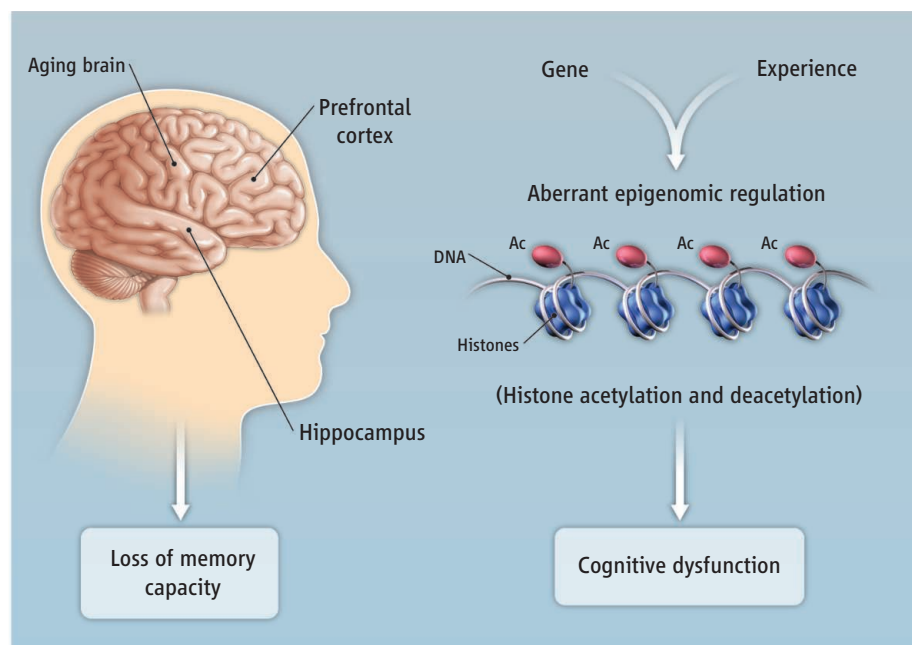
and experience-dependent epigenetic modification at the histone H4 lysine 12 (H4K12) acetylation site. This correlated with the loss of almost all normal memory-associated transcription in the hippocampus. Moreover, the authors identified a memory-associated gene, Formin 2 (an actin regulatory protein), and showed that its function is necessary for normal memory, and that its transcriptional regulation is disrupted in aging.

In a final series of studies with poten-

Changes in the epigenetic modification of chromatin may be the molecular basis for memory decline in aging adults.

tial clinical relevance, Peleg *et al.* show that intrahippocampal infusion of mice with suberoylanilide hydroxamic acid, an inhibitor of histone deacetylase, increased memory-associated H4K12 acetylation in the central nervous system, restored memory-associated transcriptional regulation, and improved behavioral memory function in aged animals.

The study presents a major advance in thinking about the role of histone modifi-



Aging and memory. Experience-epigenome interactions may drive memory formation. Decline in this system is a hypothetical basis for cognitive aging.

cations in synaptic plasticity and memory formation, and ties together three different scientific areas: chromatin regulation, memory-associated transcriptional regulation, and the molecular basis of aging-related cognitive decline. But one cautionary note in considering the work of Peleg *et al.* is to not attribute all of the memory disruption and pharmacological rescue effects in the aged animals to a single histone modification, H4K12 acetylation. Alterations in a large number of chromatin-modifying events likely occur throughout the central nervous system in aging, and improvement of memory as a result of histone deacetylase inhibition is probably due to action at multiple acetylated histone sites. Also, given the possibility of a relevant “histone code” for memory (9), even the H4K12 alterations could be tied to another epigenetic mark that is more proximally involved in the aging-associated transcriptional alterations.

There is an emerging understanding that chromatin is dynamic and is subject to extensive experience- and age-associated remodeling (7–15). For example, global loss of DNA methylation in aging, or the hypermethylation of regulatory regions (promoters) of genes associated with accelerated aging, such as the Werner syndrome and lamin A/C genes, has been proposed to control aging and longevity (13). In addition,

the sirtuins, a family of nicotinamide adenine dinucleotide (NAD)-dependent histone deacetylases, link chromatin regulation, cellular transformation, and longevity (14). And chromatin modifications also regulate telomere-length control, an aging mechanism (15). These disparate findings suggest a unifying hypothesis: that the accumulation of aberrant epigenetic marks over the life span drives aging-related cellular and physiological changes.

These considerations have led to a new hypothesis that dysregulation of epigenetic control mechanisms and the accumulation of aberrant epigenetic marks underlie aging-related cognitive dysfunction (4, 8) (see the figure). Specifically, the decreased transcription of key memory-promoting genes during aging is thought to arise from aberrant epigenetic marks and control mechanisms within brain regions particularly vulnerable to the aging process (hippocampus and prefrontal cortex), thus resulting in cognitive deficits. Further pursuit of this unifying hypothesis will require investigating the role of epigenetic molecular mechanisms that control memory formation in aging at two critical loci: histone posttranslational modifications and DNA methylation.

The work of Peleg *et al.* and others (7–9) constitutes an initial test of the capacity of manipulating the epigenome to potentially

reverse aging-associated memory dysfunction, and provide important proof-of-principle studies for evaluating whether this might be a viable approach to therapeutic intervention in cognitive aging. These studies will hopefully lead to more effective prevention strategies to improve quality of life in the aged, as well as contribute to a better understanding of memory function.

References

1. M. S. Albert, R. K. Heaton, in *Geriatric Neuropsychology*, M. S. Albert, M. B. Moss, Eds. (Guilford, New York, 1988), pp. 13–32.
2. S. N. Burke, C. A. Barnes, *Nat. Rev. Neurosci.* **7**, 30 (2006).
3. B. A. Yankner, T. Lu, P. Loerch, *Annu. Rev. Pathol.* **3**, 41 (2008).
4. S. Peleg *et al.*, *Science* **328**, 753 (2010).
5. M. W. Swank, J. D. Sweatt, *J. Neurosci.* **21**, 3383 (2001).
6. J. M. Levenson *et al.*, *J. Biol. Chem.* **279**, 40545 (2004).
7. T. L. Roth, J. D. Sweatt, *Curr. Opin. Neurobiol.* **19**, 336 (2009).
8. M. R. Penner, T. L. Roth, C. A. Barnes, J. D. Sweatt, *Front. Aging Neurosci.* **2**, 9 (2010).
9. E. Borrelli, E. J. Nestler, C. D. Allis, P. Sassone-Corsi, *Neuron* **60**, 961 (2008).
10. D. Bandyopadhyay, E. E. Medrano, *Exp. Gerontol.* **38**, 1299 (2003).
11. J. M. Sedivy, G. Banumathy, P. D. Adams, *Exp. Cell Res.* **314**, 1909 (2008).
12. C. L. Ramirez, J. Cadiñanos, I. Varela, J. M. Freije, C. López-Otin, *Cell. Mol. Life Sci.* **64**, 155 (2007).
13. M. F. Fraga, M. Esteller, *Trends Genet.* **23**, 413 (2007).
14. V. D. Longo, B. K. Kennedy, *Cell* **126**, 257 (2006).
15. M. A. Blasco, *Nat. Rev. Genet.* **8**, 299 (2007).

10.1126/science.1189968

ENGINEERING

Sewage Treatment with Anammox

B. Kartal,¹ J. G. Kuenen,² M. C. M. van Loosdrecht²

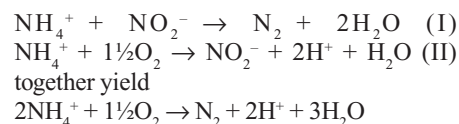
Organic matter must be removed from sewage to protect the quality of the water bodies that it is discharged to. Most current sewage treatment plants are aimed at removing organic matter only. They are energy-inefficient, whereas potentially the organic matter could be regarded as a source of energy. However, organic carbon is not the only pollutant in sewage: Fixed nitrogen such as ammonium (NH₄⁺) and nitrate (NO₃⁻) must be removed to avoid toxic algal blooms in the environment. Conventional wastewater treatment systems for nitrogen removal require a lot of energy to create aerobic conditions for bacterial nitrification, and also use organic carbon to help remove

nitrate by bacterial denitrification (see the figure). An alternative approach is the use of anoxic ammonium-oxidizing (anammox) bacteria, which require less energy (1) but grow relatively slowly. We explore process innovations that can speed up the anammox process and use all organic matter as much as possible for energy generation.

The anammox process is responsible for at least 50% of the nitrogen turnover in marine environments (2, 3) and occurs in nature at both low and high temperatures and salinities. It is a shortcut in the nitrogen cycle (see the figure) that was discovered in the early 1990s (4). The anammox bacteria, which belong to the group Planctomycetes, contain a membrane-bound organelle in which ammonium and nitrite are converted to nitrogen gas via the toxic and extremely energy-rich hydrazine intermediate. Special

Wastewater treatment including high rate anammox processes have the potential to become energy-neutral or even energy-producing.

lipids found in these bacteria, ladderanes, are believed to assist in keeping the hydrazine within this organelle (5). The bacteria use CO₂ as their carbon source for growth and hence do not require organic carbon (1). The nitrite required for their growth may be provided by aerobic ammonium-oxidizing bacteria or archaea (2). The anammox (I) and nitrification (II) reactions



In conventional sewage treatment, organic matter is combusted to carbon dioxide by microorganisms growing in flocs, generally referred to as an “activated sludge.” This process requires a lot of electrical energy input

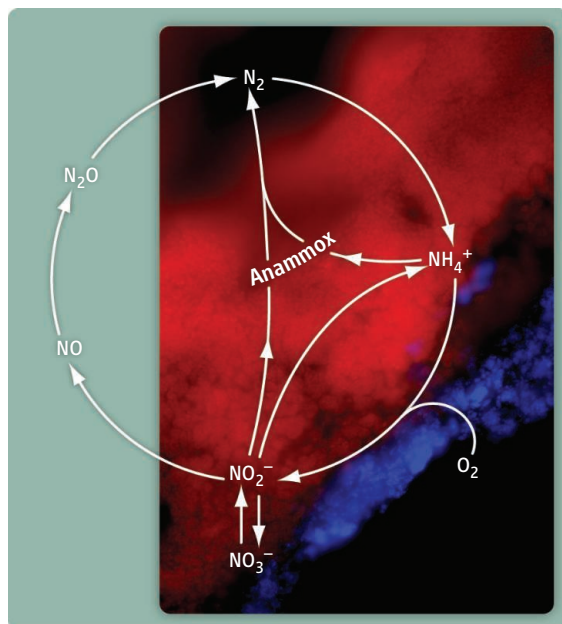
¹Microbiology, Radboud University Nijmegen, Toernooiveld 1, 6525ED Nijmegen, Netherlands. ²Biotechnology, Delft University Technology, Julianalaan 67, 2628BC Delft, Netherlands. E-mail: j.g.kuenen@tudelft.nl

for air pumps to provide oxygen for oxidation of ammonium to nitrate (see table S1 in the Supporting Online Material). Furthermore, denitrification often requires extra organic matter (methanol) for the removal of nitrate as nitrogen gas in the final step. Sometimes, a fraction of the total energy used in existing sewage treatment plants is recovered as biogas (CH_4), produced by the excess bacterial biomass (sludge).

A better approach would be to design the treatment process such that aeration is largely circumvented and most organic compounds are converted to biogas in an anaerobic treatment process. This is feasible with the present state of the art. However, in such a design, the organic compounds would no longer be available for denitrification, resulting in a large excess of ammonium and hence a potential nitrogen pollution problem.

Use of the anammox process can overcome this problem, because it does not require organic compounds for nitrogen removal (6–9). However, anammox bacteria grow slowly (generation times of 10 to 12 days at 35°C), and the anammox process is therefore currently limited to treating warm wastewater with high ammonium content. In the first designs, reactions I and II were carried out in consecutive reactors, but these were later combined in a single oxygen-limited reactor (10) where anammox and nitrite-producing bacteria coexist (see the figure). Because of the low specific conversion rates of the one-reactor process, the bottleneck in this combination has been insufficient biomass retention.

Much more biomass could be retained if fast-settling, compact granules of a coculture of nitrification and anammox bacteria (a “granular sludge”) were used. In the 1980s, the use of anaerobic methane-producing granular sludge was the key in the development of efficient anaerobic wastewater treatment in sludge blanket and fluidized bed reactors (11, 12). Granular-sludge reactors have now been developed for removal of organic matter and nutrients under aerobic conditions (13). Granular-sludge reactors achieve a very high volumetric conversion rate due to a large surface area for mass transfer. They also do not trap the inert particles that are inevitably present in wastewater and that would strongly reduce the specific activity of the sludge.



The selective production of granules has also been successfully applied on nitrifying/anammox sludge in sludge blanket reactors that are operated with oxygen-limited aeration. The result is a substantial improvement in the energy management of wastewater facilities, as reported for a case in the Netherlands (14). Because of their high volumetric conversion rate, granular-sludge systems also offer the possibility for application of anammox for sewage treatment at the low temperatures and concentrations that are typical for municipal wastewater.

The availability of a fast anammox process opens real perspectives for a complete redesign of the present energy-consuming into an energy-yielding wastewater treatment (see table S1). The principle of this process would be as follows. Sewage is first led to a very-high-load activated sludge system, where soluble organic matter is converted to biomass with maximal growth yield that can be flocculated and separated together with the nondegraded suspended and colloidal material in a settler. In this way most organic matter is removed and concentrated, enabling its use in the generation of biogas (methane) in a digestion process. The effluent of the first stage is combined with the digester effluent and treated in a granular-sludge anammox reactor. The remaining nitrate will be below the required effluent standards for nitrogen discharge to the environment.

Calculations based on (7) show that the wastewater treatment process with anammox in the main line would yield 24 watt hours per person per day ($\text{Wh p}^{-1} \text{d}^{-1}$), compared to a consumption of $44 \text{ Wh p}^{-1} \text{d}^{-1}$ in conventional treatment. The realization of such a

The nitrogen cycle including anammox. Molecular nitrogen (N_2) is fixed biologically or industrially to ammonium (NH_4^+), the main fertilizer for plants. When ammonium is released to the environment, it may be oxidized by aerobic, nitrifying bacteria and archaea to nitrite (NO_2^-) and nitrate (NO_3^-), respectively, which plants can use as an additional nitrogen source. Under anaerobic conditions, nitrate and nitrite may be reduced back to ammonium, or to nitrogen gas through denitrification. Nitrite can also be combined with ammonium to give nitrogen gas in the anammox reaction. **(Background)** Outer layer of a compact nitrogen-producing granule for possible use in energy-generating wastewater treatment. The anammox bacteria (red) are on the inside of the granule; the nitrite-producing bacteria (blue) reside in a $40\text{-}\mu\text{m}$ -thick layer on the outside, ensuring that oxygen does not reach the anoxic anammox bacteria. The bacteria have been stained with fluorescent 16S rDNA probes.

substantial gain presents great challenges in terms of technology and investment. However, we are convinced that it is feasible and expect that these and other innovations may stimulate investigators to take up the challenge of making global wastewater treatment energy-neutral or even energy-generating, as called for by many governmental programs, such as (15).

The discovery and practical application of high-rate nitrogen removal processes for wastewater treatment offer an enormous opportunity to make our wastewater treatment not only sustainable, but also a basis for the production of clean water that can be recycled for a variety of purposes.

References and Notes

1. J. G. Kuenen, *Nat. Rev. Microbiol.* **6**, 320 (2008).
2. C. A. Francis, J. M. Beman, M. M. M. Kuypers, *ISME J.* **1**, 19 (2007).
3. K. R. Arrigo, *Nature* **438**, 122 (2005).
4. A. Mulder et al., *FEMS Microbiol. Ecol.* **16**, 177 (1995).
5. J. S. Sinninghe Damste et al., *Nature* **419**, 708 (2002).
6. M. S. M. Jetten, S. J. Horn, M. C. M. van Loosdrecht, *Water Sci. Technol.* **35**, 171 (1997).
7. H. Siegrist, D. Salzgeber, J. Eugster, A. Joss, *Water Sci. Technol.* **57**, 383 (2008).
8. W. R. van der Star et al., *Water Res.* **41**, 4149 (2007).
9. B. Wett, *Water Sci. Technol.* **56**, 81 (2007).
10. K. A. Third, A. O. Sliemers, J. G. Kuenen, M. S. M. Jetten, *Syst. Appl. Microbiol.* **24**, 588 (2001).
11. G. Lettinga, L. W. H. Pol, *Water Sci. Technol.* **24**, 87 (1991).
12. G. Lettinga, A. F. M. Vanvelsen, S. W. Hobma, W. Dezeuw, A. Klapwijk, *Biotechnol. Bioeng.* **22**, 699 (1980).
13. M. K. de Kreuk, M. Pronk, M. C. M. van Loosdrecht, *Water Res.* **39**, 4476 (2005).
14. W. R. Abma, M. C. M. van Loosdrecht, *Water Sci. Technol.* **61**, 1715 (2010).
15. See energy.ca.gov/water-energy/index.html or epa.gov/waterinfrastructure/bettermanagement_energy.html

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/702/DC1
Table S1

10.1126/science.1185941

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/704>

This article **cites 33 articles**, 14 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/328/5979/704#otherarticles>

This article appears in the following **subject collections**:

Oceanography

<http://www.sciencemag.org/cgi/collection/oceans>

Bioluminescence in the Ocean: Origins of Biological, Chemical, and Ecological Diversity

E. A. Widder

From bacteria to fish, a remarkable variety of marine life depends on bioluminescence (the chemical generation of light) for finding food, attracting mates, and evading predators. Disparate biochemical systems and diverse phylogenetic distribution patterns of light-emitting organisms highlight the ecological benefits of bioluminescence, with biochemical and genetic analyses providing new insights into the mechanisms of its evolution. The origins and functions of some bioluminescent systems, however, remain obscure. Here, I review recent advances in understanding bioluminescence in the ocean and highlight future research efforts that will unite molecular details with ecological and evolutionary relationships.

The vast majority of bioluminescent organisms reside in the ocean; of the more than 700 genera known to contain luminous species, some 80% are marine (1). These occupy a diverse range of habitats, from polar to tropical and from surface waters to the sea floor (2). The ecological importance of bioluminescence in the ocean is manifest in the dominance of light emitters in open waters; luminescent fish (e.g., myxophids and hatchetfish) and crustaceans (e.g., copepods, krill, and decapods) dominate in terms of biomass, whereas bacteria and dinoflagellates dominate in terms of abundance (3, 4). Its import is also evident in the large number of organisms that retain functional eyes to detect bioluminescence at depths where sunlight never penetrates and in the remarkable degree of diversity and evolutionary convergence among light-emitting organisms (4).

Bioluminescent species are found in most of the major marine phyla from bacteria to fish. As a phylum, comb jellies have the highest proportion of bioluminescent species, whereas other phyla such as diatoms and arrow worms have none or few luminescent representatives (2, 4).

Rivaling its diverse distribution is its impressive array of colors, intensities, and kinetics. Measurements of bioluminescent emission spectra have revealed a rainbow palette of hues that extend over the full visible range (Fig. 1) (5–8). Because most bioluminescence has evolved in the open ocean, most emission spectra are blue, centered on the wavelength that travels farthest through seawater ($\lambda_{\text{max}} \sim 475 \text{ nm}$) (4). Green is the next most common color and is more often found in benthic and shallow coastal species, possibly because increased turbidity from particles in the water scatters blue light and favors the transmission of longer wavelengths (6, 9). Violet, yellow, orange, and red occur only rarely, and in most of

these cases their functions and chemistries remain obscure (1, 5–8).

Photon fluxes span at least nine orders of magnitude, from about 10^3 photons per second for a single bioluminescent bacterium to more than 10^{12} photons per second for some krill and fish (10–12). Emission kinetics range from the persistent glow of bioluminescent bacteria to flashes as brief as 43 ms from lanternfish light organs (10). Luminescent chemicals may be released directly into the water or retained within cells called photocytes. The angular distribution and waveband of light emitted by photocytes may be adjusted by means of muscles and complex optical components that reflect, refract, or filter the light, in which case the photocytes and accessory structures are called photophores or light organs. Emitters may also produce spatial patterns of light displayed over the surface of their bodies or by swimming patterns during light emission (4). All of these parameters carry information to the eyes of potential predators, prey, or members of the same species.

Understanding what function bioluminescence serves in a particular organism provides insight into what selection pressures imposed by the environment and by intergroup competition may have favored the evolution of bioluminescence in one group over another. Wide diversity among light-emitting chemistries has long confounded efforts to trace evolutionary origins. Here, I review new evidence centered on alternative cellular functions for light-emitting molecules and genomic analyses of light emitters that further illuminates the evolutionary origins of bioluminescence.

What Are the Functions of Bioluminescence?

The many functions of bioluminescence reflect the unique nature of the visual environment in which they have evolved. The open ocean is a world without hiding places, where sunlight filtering down through clear water diminishes approximately 10-fold for every 75 m of descent, until all

visible light disappears below 1000 m (12). Under sunlight or moonlight, the light field is dim, blue, and highly directional. In order to hide, many animals vertically migrate downward into the dark depths during the day and only venture into food-rich surface waters under cover of darkness (4). This results in what some consider the most massive animal migration pattern on the planet (12). As a consequence of this migration, most open ocean inhabitants live their lives in dim light or darkness, where bioluminescence can aid animal survival in at least three critical ways: (i) It can serve as an aid in locating food, either by means of built-in headlights or by the use of glowing lures. (ii) It can be used to attract a mate by means of species-specific spatial or temporal patterns of light emission. (iii) It can function as a defense against predators (4). The last is probably the most common use and takes many forms. Some animals, including crustaceans, squid, jellyfish, and fish, release their light-emitting chemicals into the water, producing clouds or particles of light that serve to distract or blind a predator (2, 4, 12). Other animals mark their predators with luminescent slime, making them easy targets for secondary predators (2). Alternatively, when caught in the clutches of a predator, some luminescent prey produce bright and often elaborate displays, which attract secondary predators that will attack the first attacker, thereby affording them an opportunity for escape (2, 4). Luminescence may also be used as a warning to predators, signaling the unpalatability of the prey (2). It is also used extensively as camouflage, in a process called counterillumination, whereby the silhouette of an opaque animal is replaced by bioluminescence of comparable color, intensity, and angular distribution to downwelling ambient light. This latter use of bioluminescence is common among fishes, crustaceans, and squid that inhabit the twilight depths of the ocean where many predators have upward-looking eyes adapted for locating the silhouettes of prey (2, 4, 9).

In most cases, the presumed function of the light emission has its basis in inference from morphological and physiological characteristics rather than experimental studies or in situ observations (13). For example, in the case of fish with red-emitting light organs (Fig. 1), the location of the light organ just below the eye and the unusual long-wavelength sensitivity of the eye suggest that their red luminescence may be used to illuminate prey that are blind to red light (14). In the case of luminous bacteria that form specific symbioses with certain marine fishes and squid, the adaptive value of the light emission is generally evident: The bacteria provide the host with light that can be used to attract prey, evade predators, or attract a mate, while the host provides the bacteria with an ideal growth environment (15). For free-living bacteria where the adaptive value is less evident, the most generally accepted hypothesis is that luminous bacteria growing on fecal pellets may serve as

Ocean Research and Conservation Association, Fort Pierce, FL 34949, USA. E-mail: ewidder@teamorca.org

an attractant, causing the pellets to be consumed and thereby introducing the bacteria to an animal's nutrient-rich gut (2).

Experimental evidence for function has been largely confined to studies of physiological control, such as experiments demonstrating the ability of counterillumination to adjust the physical characteristics of their ventral light emissions to match those of experimentally manipulated downwelling light fields (4, 9). Behavioral experiments are far less common, with the most extensive studies being those with dinoflagellates demonstrating that their light emission reduces grazing by nocturnal predators (4).

Opportunities for direct in situ observation are rare. Explorations with submersibles and remote-operated vehicles are regularly revealing new luminescent organisms, such as the newly discovered bombardier worms: swimming deep-sea annelids that release green light bombs when disturbed (16). But many behaviors can only be observed unobtrusively by using methodologies that have recently become possible with far-red

illumination and intensified imaging technologies (17).

What Is the Biochemical Variability of Bioluminescence?

The chemical reaction involved in bioluminescence must be sufficiently energetic to produce an excited singlet state molecule that will generate a visible photon as it relaxes back down into its ground state (in contrast to fluorescence and phosphorescence, which depend on absorption of sufficiently energetic photons). Chemical oxidation reactions involving molecular oxygen fit this criterion (1), which may explain why the primary mechanism operating in bioluminescent reactions involves the breakdown of a peroxide bond (18). In fact, the generic terms for the enzyme (luciferase) and substrate (luciferin) involved in light-producing reactions require taxon prefixes to distinguish the different bioluminescent systems (Fig. 2).

In bacteria, two simple substrates [a reduced flavin mononucleotide (FMNH₂) and a long-

chain aliphatic aldehyde (RCHO)] are oxidized by molecular oxygen and luciferase. The aldehyde is consumed during the reaction but is continuously synthesized by the bacteria, resulting in a persistent glow (15). Alternatively, the chemical structure of dinoflagellate luciferin bears a striking similarity to chlorophyll (Fig. 2), which suggests that it originated in photosynthetic species. Although the biosynthetic pathway of luciferin is unknown in dinoflagellates (19), a dietary dependence on dinoflagellate luciferin has been suggested in krill (2). Ostracod luciferin is an imidazopyrazinone synthesized from three amino acids (Trp-Ile-Arg) as is coelenterazine (Phe-Tyr-Tyr) (Fig. 2), but in both cases the details of biosynthesis are unknown (2). In the case of coelenterazine, its manner of biosynthesis has recently become of particular interest with the discovery that coelenterates require it as a dietary source (20). Although there is some circumstantial evidence for its synthesis in crustaceans (21), such a linkage remains to be confirmed. In some bioluminescent systems, accessory proteins serve as secondary emitters, which shift

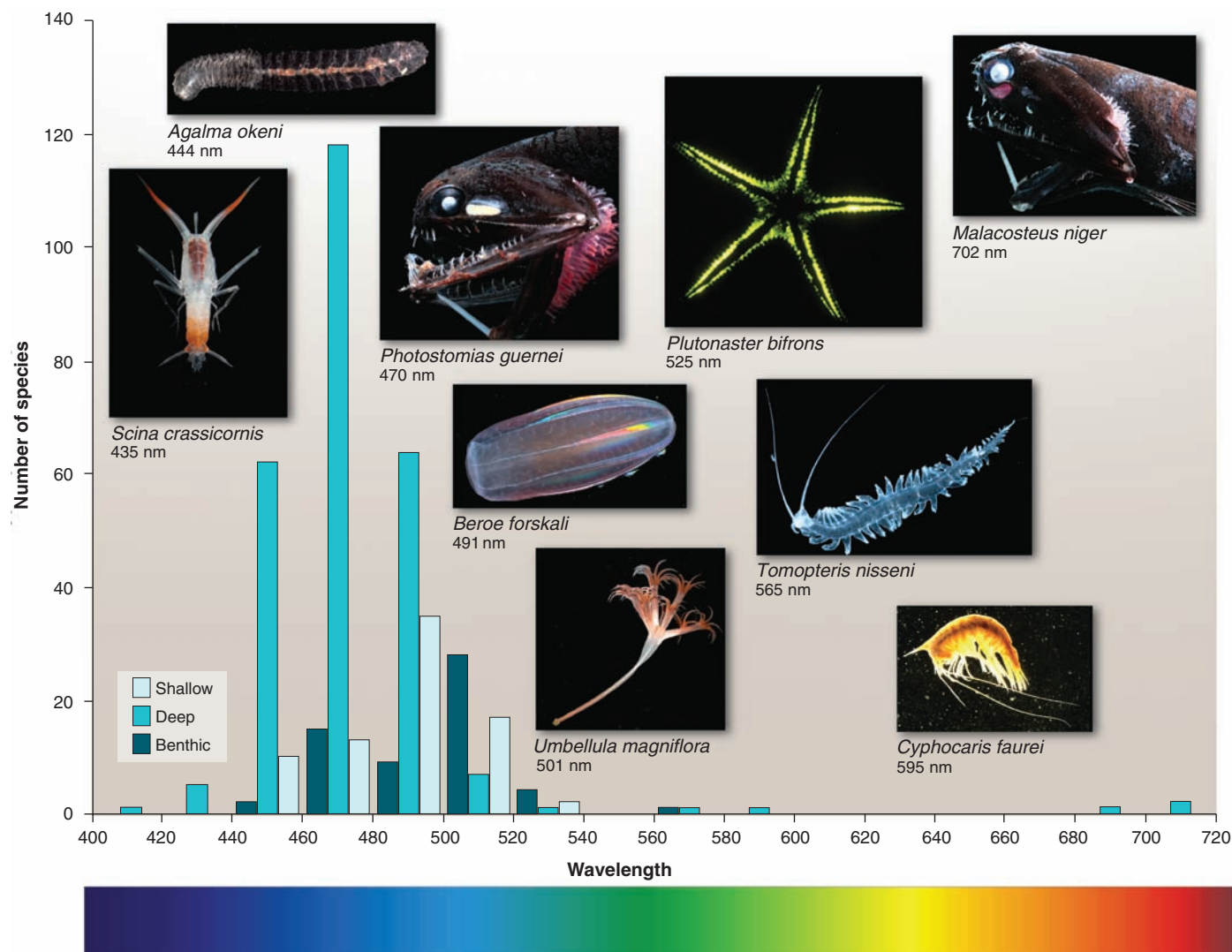


Fig. 1. The distribution of bioluminescence emission maxima varies by marine environment and organism type. Bioluminescent emissions extend over the full visible range and beyond. [Photo credits: J. Cohen for the photograph of *S. crassicornis*; P. Herring, *P. bifrons*; and P. Batson (DeepSeaPhotography.com), *C. faurei*]

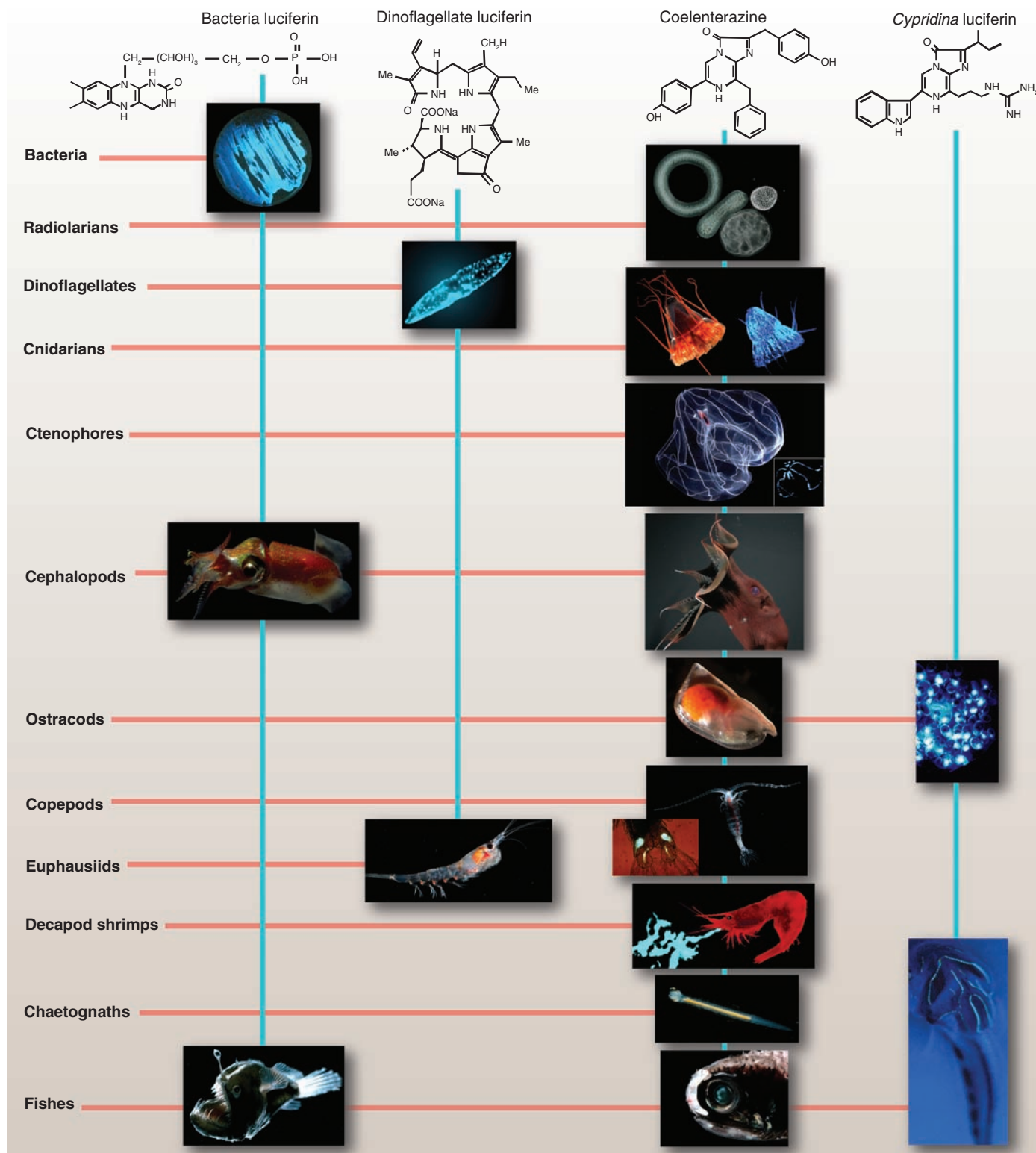


Fig. 2. The chemical structures of the four best-known luciferins are as diverse as their phylogenetic distribution. Bacterial luciferin may occur in free-living or symbiont bacteria (e.g., in squid such as *Heteroteuthis dispar*) or in fish such as *Melanocetus johnsoni*. Dinoflagellate luciferin occurs not only in dinoflagellates (e.g., *Pyrocystis fusiformis*) but also in euphausiids (e.g., *Meganctiphanes norvegica*). Some of those using coelenterazine as luciferin include radiolarians (e.g., unidentified polycystine radiolarians), cnidarians (e.g., scyphozoan *Periphylla periphylla*, as seen in the light and photographed by its own light), ctenophores (e.g., *Bathocyroe fosteri*, with bioluminescence display shown in inset), vampire squid (e.g., *Vampyroteuthis*

infernalis), ostracods (e.g., *Orthoconchoecia agassizi*), copepods (e.g., *Gaussia princeps* releasing its bioluminescent chemicals from glands on its tail, shown in inset), decapods (e.g., *Acanthephyra purpurea* spewing luciferin and luciferase out of its mouth), chaetognaths (e.g., *Caecosagitta macrocephala*), and fish (e.g., the myctophid *Diaphus* sp. has a large preorbital light organ). Cypridina luciferin, which is an imidazopyrazinone like coelenterazine, is found in ostracods such as *Vargula hilgendorffii* and is the dietary source of luciferin for the midshipman fish *Porichthys notatus*. [Photo credits: S. Haddock, radiolarians and chaetognath; K. Reisenbichler, *V. infernalis*; J. Case, copepod luminescent glands and midshipman fish photophores]

the color of the bioluminescent emission to longer wavelengths. The best known of these is green fluorescent protein (GFP), which was isolated and cloned from a bioluminescent jellyfish and has been used extensively as an *in vivo* fluorescent marker of gene expression, protein synthesis, and cell lineage (1).

Besides the four best-known luciferins used by marine organisms (Fig. 2), there are some that are partially defined, such as that of the rock-boring clam *Pholas dactylus*, the parchment tube worm *Chaetopterus variopedatus*, and the syllid fireworm *Odontosyllis enopla*, as well as many more yet to be elucidated, most notably among mollusks, echinoderms, and hemicordates (acorn worms) (1).

What Are Evolutionary Processes That Lead to Bioluminescence?

Based on the number of light-producing chemistries across the monophyletic lineages, bioluminescence is estimated to have evolved independently at least 40 times (2). Remarkably, not only is there evidence of independent origins within taxa (e.g., ostracods have two known chemistries: coelenterazine and vargulin) but even within individual species (e.g., the deep-sea anglerfish, *Linophryne coronata*, has two different light-emitting systems in adult females: bacterial luminescence in the dorsal lure and an intrinsic, unidentified chemistry in the chin barbel) (Fig. 3A).

Most hypotheses put forth to explain the evolution of luminescent systems fall into two basic categories related to selection acting on either substrates or enzymes. In the first case, selection driving the evolution of luciferin substrates may have resulted from pressures to protect organisms from photochemically generated reactive oxygen species such as H_2O_2 and O_2^- (18). For example, the luciferin coelenterazine, which is found in at least nine phyla, is a strong antioxidant (2, 18). As vision-dependent animals migrated to greater depths to escape detection by visual predators, the reduced oxidative stress in deeper waters shifted the selection pressure from the antioxidative to the chemiluminescent properties of this molecule (18).

The alternative enzyme-centric explanation suggests luciferases originally acted as mixed-function oxygenases (22). In this case, as visual animals were driven into darker waters, natural selection may have favored the development of more-sensitive eyes and enhanced visibility of visual signals (4). As a consequence, a mutation in an oxygenase enzyme involved in the breakdown of pigment molecules associated with spots displayed to attract a mate or repel a predator could result in external luminescence that caused immediate selective pressures for the light emitter (22). Although the oxygenase hypothesis has been questioned by genetic and biochemical evidence (18), there remains support for enhanced visual signaling selection pressures. For example, there is evidence that the deep-sea finned octopod, *Stauroteuthis syrtensis*, developed light organs from suckers because the selective advantage of

the visual display of suckers (i.e., to attract a mate) superseded the advantage provided by their adhesive properties (Fig. 3B) (23).

Although the chemical basis for light production in octopods has not been determined, light emission at suckers or any other localized photophore would require an enzyme mutation coupled with substrate availability. Therefore, by combining the substrate-centric and enzyme-centric hypotheses, one might envision a scenario where reduced oxidative stress in deeper waters freed up antioxidants, such as coelenterazine, as substrates for chemiluminescent reactions that resulted from specific enzyme mutations. This view gains support from the fact that, although light-emitting substrates are relatively few in number and conserved across phyla, the bioluminescent enzymes are unique and independently derived (2).

There may also be a connection to protection from oxidative stress with GFP because similar proteins exist in bioluminescent jellyfish and sea pens, as well as in non-bioluminescent corals, copepods, and lancelets (24). In corals and lancelets, it has been suggested that GFP-like proteins could function as antioxidants to detoxify reactive oxygen species (24, 25).

In bioluminescent bacteria, the question of evolutionary origins has recently gained new focus with the reclassification of members of the *Vibrio fischeri* species group as a new genus, *Aliivibrio* (26). The taxonomy of luminescent bacteria has been revised often in efforts to better define evolutionary relationships and origins. The distribution of bioluminescent species among bacteria is not even; all species in the terrestrial genus *Photorhabdus* are luminescent, but marine genera with bioluminescent species (*Aliivibrio*, *Photobacterium*, *Shewanella*, and *Vibrio*) include many closely related nonluminescent species (15). Nonetheless, comparative genomics have revealed that all luminous bacteria share a common gene sequence: the *lux* operon that encodes for the biosynthesis of luciferase and its substrates (15). This highly conserved sequence appears in bacteria from very different ecological niches, suggesting a strong selective advantage despite the energetic costs of producing light. In mixed cultures of luminescent and dark mutants of *Vibrio harveyi*, the dark mutants rapidly overrun the culture unless the mixture is irradiated with ultraviolet (UV) light, in which case the balance tips the other way, apparently because bioluminescence stimulates DNA repair (27). If DNA repair was the initial selective advantage for light production in bacteria, then the *lux* operon may have been lost in bacteria that evolved more efficient DNA repair systems but retained in those where visible light became a selective advantage. Further selective advantage would have been afforded with the evolution of quorum sensing, which conserves energy by assuring that luminescent bacteria do not synthesize their light-producing chemicals unless a sufficient concentration are present to be visible. Although once considered confined to bioluminescent bacteria, quorum sensing is widespread in nonluminescent

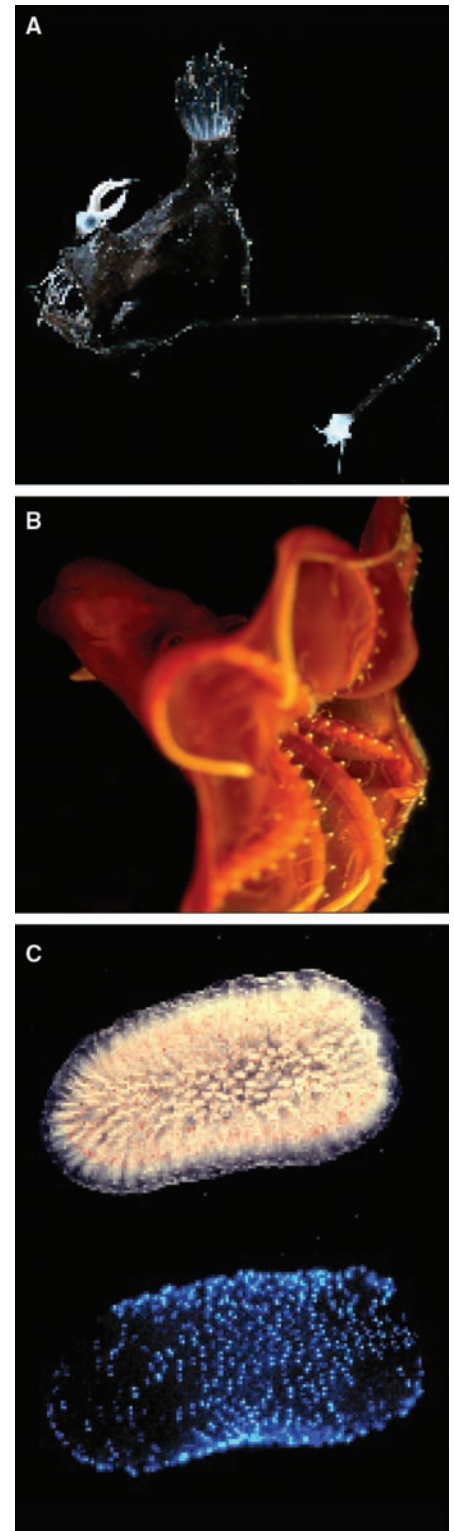


Fig. 3. Bioluminescence has resulted from some intriguing evolutionary adaptations. (A) In the deep-sea anglerfish *Linophryne coronata*, bioluminescence from the esca is bacterial in origin, whereas that from the chin barbel is an unidentified intrinsic chemistry. (B) In the octopus *Stauroteuthis syrtensis*, its suckers are photophores. (C) In the tunicate *Pyrosoma atlanticum*, luminescence originates from putative bacterial endosymbionts.

Gram-negative bacteria, where it serves multiple functions, such as enhancing pathogenicity of bacteria by delaying toxin production until population densities are high enough to overwhelm the host's defenses (28). The question, therefore, arises: Which came first, quorum sensing or bioluminescence?

The details of bacterium-host interactions in bioluminescent symbionts [e.g., *A. fischeri* and the Hawaiian bobtail squid *Euprymna scolopes* (29–31)] reveal unique insights into the co-evolution of bioluminescence among distinct classes of organisms. For example, the recent discovery of light perception capabilities within the light organ of *E. scolopes* explains how the host has the means to reject nonluminescent strains of *A. fischeri* (32). Although it was once thought that such complex and tightly coupled associations must have coevolved, recent phylogenetic analyses of bacteria isolated from two squid families (33) and seven teleost families (34) revealed deep divergences among the hosts that are not reflected in the symbionts, pointing to evolutionarily independent origins of these symbioses. It will be interesting to see whether the same differences exist between hosts and symbionts in flashlight fishes and deep-sea anglerfishes, where the bacterial symbionts, which are re-acquired from the environment with each generation, are as yet unidentified and unculturable (15).

Further along the continuum toward greater integration of symbiont and host are the colonial tunicates known as pyrosomes (Fig. 3C). On the basis of demonstrated bacterial luciferase activity, their luminescent organs may be bacterial endosymbionts (1). However, they do not produce the persistent glow characteristic of luminous bacteria but emit light in response to either mechanical or photic stimuli (12). Pyrosomes are exciting candidates for genomic-level analysis and possibly gaining new insight into the mechanisms that underlie endosymbiosis.

Some of the most detailed information on the evolution of bioluminescent enzymes in the ocean comes from work on marine dinoflagellates. Although there are no sequenced dinoflagellate genomes, eight dinoflagellate luciferase genes have been fully identified (19). Phylogenetic analysis indicates that the most primitive luciferase of these eight is from the large heterotroph *Noctiluca scintillans*. This luciferase gene codes for a single catalytic domain (19), whereas in the seven other sequenced luciferases—all from photosynthetic dinoflagellates—there are three homologous catalytic sites (35). Although these sites are highly conserved across this group, there are some differences in their genetic structure; for example, the luciferase gene in *Pyrocystis* has a large unique noncoding region (36). *P. lumula* also differs substantially from the other dinoflagellates in this group in that it lacks luciferin

binding protein and exhibits no circadian rhythm in the breakdown of luciferase, luciferin binding protein, and its light-producing organelles and it produces approximately two orders of magnitude more light in response to mechanical stimuli (12). Given these differences, it seems likely that the selective pressures shaping the luminescent capacities of these dissimilar dinoflagellates may have diverged at some point in evolutionary history. Relating comparative genomic analysis to variable luminescent capacities and control mechanisms in different dinoflagellates is an intriguing new approach to comprehending the adaptive importance of bioluminescence. Dinoflagellates may also provide valuable insight into how gene duplication and gene loss function in generating biodiversity. For example, luciferase gene loss can now be examined by using oligonucleotide primers recently developed for specific luciferase genes. Use of these primers revealed a strain of *Gonyaulax spinifera* that produces bioluminescence at such low levels it is undetectable to the human eye, suggesting that either the luciferase is not expressed or that only remnants of the genes remain (37).

Seeing the Light

The many examples of evolutionary convergence related to bioluminescence are a testament to the survival value of the trait, whereas its abundance and ubiquity in the ocean attests to its importance in marine ecosystems. In situ imaging systems used to document the vertical distribution of planktonic emitters from surface to sea floor (38), as well as calculate their nearest-neighbor distances (12), will become increasingly important to providing a more detailed understanding of animal distributions and population dynamics in ocean ecosystems. Satellite sensor systems may also contribute, as with the recent detection of a 15,400-km² bioluminescent “milky sea” in the Indian Ocean (39). This eerie phenomenon may be due to a luminous bacterium (e.g., *V. harveyi*) growing on the remains of a monsoon-induced algal bloom (12). However, this hypothesis awaits confirmation until in situ collections are made in an active milky sea. Improved low-light sensors deployed from aerial platforms may facilitate more targeted sampling efforts. It is hoped that with improved in situ sensor technology and additional observation platforms, such as autonomous underwater vehicles and undersea observatories, new insights from the field will be combined with more detailed genomic and physiologic studies in the laboratory to better understand the ecological importance and adaptive value of bioluminescence in the ocean.

References and Notes

1. O. Shimomura, *Bioluminescence: Chemical Principles and Methods* (World Scientific, Toh Tuck Link, Singapore, 2006).

2. S. H. D. Haddock, M. A. Moline, J. F. Case, *Annu. Rev. Mar. Sci.* **2**, 443 (2010).
3. P. J. Herring, E. A. Widder, in *Encyclopedia of Ocean Science* Vol. 1, J. H. Steele, S. A. Thorpe, K. K. Turekian, Eds. (Academic Press, San Diego, CA, 2001), pp. 308–317.
4. E. A. Widder, in *Adaptive Mechanisms in the Ecology of Vision*, S. N. Archer, M. B. A. Djamgoz, E. Loew, J. C. Partridge, S. Vallegria, Eds. (Kluwer Academic, Dordrecht, Netherlands, 1999), pp. 555–581.
5. S. H. D. Haddock, J. F. Case, *Mar. Biol.* **133**, 571 (1999).
6. P. J. Herring, *Proc. R. Soc. Lond.* **220**, 183 (1983).
7. M. I. Latz, T. M. Frank, J. F. Case, *Mar. Biol.* **98**, 441 (1988).
8. E. A. Widder, M. I. Latz, J. F. Case, *Biol. Bull.* **165**, 791 (1983).
9. S. Johnsen, E. A. Widder, C. D. Mobley, *Biol. Bull.* **207**, 1 (2004).
10. A. F. Mensinger, J. F. Case, *J. Exp. Mar. Biol. Ecol.* **210**, 75 (1997).
11. S. D. Miller, S. H. D. Haddock, C. D. Elvidge, T. F. Lee, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 14181 (2005).
12. E. A. Widder, *Mar. Freshwat. Behav. Physiol.* **35**, 1 (2002).
13. P. J. Herring, *J. Mar. Biol. Assoc. U.K.* **87**, 829 (2007).
14. E. A. Widder, M. I. Latz, P. J. Herring, J. F. Case, *Science* **225**, 512 (1984).
15. P. V. Dunlap, in *Encyclopedia of Microbiology*, M. Schaechter, Ed. (Elsevier, Oxford, 2009), pp. 45–61.
16. K. J. Osborn, S. H. D. Haddock, F. Pleijel, L. P. Madin, G. W. Rouse, *Science* **325**, 964 (2009).
17. E. A. Widder, *Oceanography* **20**, 46 (2007).
18. J. F. Rees et al., *J. Exp. Biol.* **201**, 1211 (1998).
19. L. Liu, J. W. Hastings, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 696 (2007).
20. S. H. D. Haddock, T. J. Rivers, B. H. Robison, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11148 (2001).
21. C. M. Thomson, P. J. Herring, A. K. Campbell, *J. Mar. Biol. Assoc. U.K.* **75**, 165 (1995).
22. H. H. Seliger, *Nav. Res. Rev.* **XLV**, 5 (1993).
23. S. Johnsen, E. J. Balser, E. A. Widder, *Nature* **398**, 113 (1999).
24. D. D. Deheyn et al., *Biol. Bull.* **213**, 95 (2007).
25. C. H. Mazel et al., *Limnol. Oceanogr.* **48**, 402 (2003).
26. H. Urbanczyk, J. C. Ast, M. J. Higgins, J. Carson, P. V. Dunlap, *Int. J. Syst. Evol. Microbiol.* **57**, 2823 (2007).
27. A. Czyz, K. Plata, G. Wegrzyn, *Luminescence* **18**, 140 (2003).
28. J. W. Hastings, E. P. Greenberg, *J. Bacteriol.* **181**, 2667 (1999).
29. M. J. McFall-Ngai, in *Gut Microbiota and Regulation of the Immune System*, G. Huffnagle, M. Noverr, Eds. (Landes Bioscience, Austin, TX, 2008), pp. 102–112.
30. S. V. Nyholm, M. J. McFall-Ngai, *Nat. Rev. Microbiol.* **2**, 632 (2004).
31. E. G. Ruby, *Nat. Rev. Microbiol.* **6**, 752 (2008).
32. D. Tong et al., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9836 (2009).
33. R. C. Guerrero-Ferreira, M. K. Nishiguchi, *Cladistics* **23**, 497 (2007).
34. P. V. Dunlap et al., *Cladistics* **23**, 507 (2007).
35. L. Liu, T. Wilson, J. W. Hastings, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 16555 (2004).
36. O. K. Okamoto, L. Liu, D. L. Robertson, J. W. Hastings, *Biochemistry* **40**, 15862 (2001).
37. A. Baker, I. Robbins, M. A. Moline, M. D. Iglesias-Rodriguez, *J. Phycol.* **44**, 419 (2008).
38. I. G. Priede, P. M. Bagley, S. Way, P. J. Herring, J. C. Partridge, *Deep-Sea Res.* **53**, 1272 (2006).
39. S. D. Miller, S. H. D. Haddock, C. D. Elvidge, T. F. Lee, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 14181 (2005).
40. I am grateful to P. Herring for helpful comments and the Office of Naval Research for its long-standing support of bioluminescence research.

10.1126/science.1174269

Washing Away Postdecisional Dissonance

Spike W. S. Lee* and Norbert Schwarz

Hand washing removes more than dirt—it also removes the guilt of past misdeeds, weakens the urge to engage in compensatory behavior (1), and attenuates the impact of disgust on moral judgment (2). These findings are usually conceptualized in terms of a purity-morality metaphor that links physical and moral cleanliness (3); however, they may also reflect that washing more generally removes traces of the past by metaphorically wiping the slate clean. If so, washing one's hands may lessen the influence of past behaviors that have no moral implications at all. We test this possibility in a choice situation. Freely choosing between two similarly attractive options (e.g., Paris or Rome for vacation) arouses cognitive dissonance, an aversive psychological state resulting from conflicting cognitions. People reduce dissonance by perceiving the chosen alternative as more attractive and the rejected alternative as less attractive after choice, thereby justifying their decision (4, 5). We test whether hand washing reduces this classic postdecisional dissonance effect.

In individual sessions as part of an alleged consumer survey, 40 undergraduates browsed 30 CD covers as if they were in a music store. They selected 10 CDs they would like to own and ranked them by preference. Later, the experimenter offered them a choice between their fifth- and sixth-ranked CDs as a token of appre-

ciation from the sponsor. After the choice, participants completed an ostensibly unrelated product survey that asked for evaluations of a liquid soap; half merely examined the bottle before answering, whereas others tested the soap by washing their hands. After a filler task, participants ranked the 10 CDs again, allegedly because the sponsor wanted to know what people think about the CDs after leaving the store (6).

Can washing one's hands attenuate the need to justify a recent choice? Yes (Fig. 1). For those who merely examined the soap, the preference for the chosen over the rejected alternative increased from before choice [mean (M) = 0.14 and SD = 1.01] to after choice (M = 2.05, SD = 1.96) by an average of 1.9 ranks [$F(1, 38)$ = 20.40, P < 0.001, for the simple effect], replicating the standard dissonance effect. In contrast, for those who washed their hands, preferences were unaffected by their decision [before choice, M = 0.68, SD = 0.75; after choice, M = 1.00, SD = 1.41; F < 1 for the simple effect]. Thus, hand washing significantly reduced the need to justify one's choice by increasing the perceived difference between alternatives [$F(1, 38)$ = 6.74, P = 0.01, for the interaction of time and hand-washing manipulation].

A study with a different choice task, cleaning manipulation, and measure replicated this finding (7). In individual sessions, 85 students responded

to an alleged consumer survey about four fruit jams (shown in pictures). They were subsequently offered a choice between two jars as a sign of the sponsor's appreciation. After their choice, participants completed an ostensibly unrelated product survey about an antiseptic wipe; half merely examined the wipe, whereas others tested it by cleaning their hands. Next, they rated the expected taste of the four jams (0 = not good at all; 10 = very good).

Participants who did not clean their hands after making a choice expected the chosen jam to taste much better (M = 8.00, SD = 1.65) than the rejected jam (M = 6.43, SD = 1.81) [$F(1, 83)$ = 27.54, P < 0.001, for the simple effect]; hand cleaning attenuated this difference to nonsignificance [for chosen jam, M = 7.63 and SD = 1.56; for rejected jam, M = 7.23 and SD = 1.25; $F(1, 83)$ = 1.79, P = 0.19, for the simple effect]. Thus, hand cleaning significantly reduced the classic postdecisional dissonance effect [$F(1, 83)$ = 7.80, P = 0.006, for the interaction of product and hand-cleaning manipulation].

These findings indicate that the psychological impact of physical cleansing extends beyond the moral domain. Much as washing can cleanse us from traces of past immoral behavior, it can also cleanse us from traces of past decisions, reducing the need to justify them. This observation is not captured by the purity-morality metaphor and highlights the need for a better understanding of the processes that mediate the psychological impact of physical cleansing. To further constrain the range of plausible candidate explanations, future research may test whether the observed "clean slate" effect is limited to past acts that may threaten one's self-view (e.g., moral transgressions and potentially poor choices) or also extends to past behaviors with positive implications.

Difference between choice alternatives

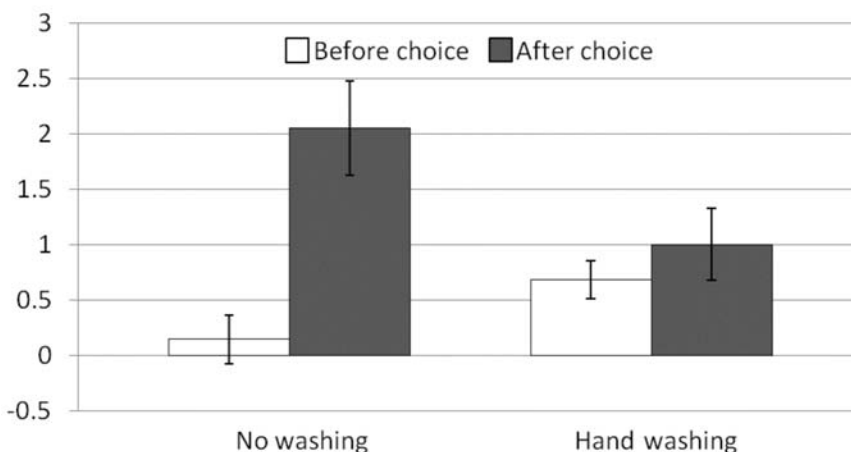


Fig. 1. Postdecisional dissonance after hand washing or no hand washing (study 1). Each bar represents the rank difference between the chosen and rejected alternatives, with higher values indicating higher preferences for the chosen alternative. Error bars represent standard error.

References and Notes

1. C.-B. Zhong, K. Liljenquist, *Science* **313**, 1451 (2006).
2. S. Schnall, J. Benton, S. Harvey, *Psychol. Sci.* **19**, 1219 (2008).
3. G. Lakoff, M. Johnson, *Philosophy in the Flesh* (Basic, New York, 1999).
4. L. Festinger, *A Theory of Cognitive Dissonance* (Stanford Univ. Press, Stanford, CA, 1957).
5. J. Cooper, *Cognitive Dissonance: 50 Years of a Classic Theory* (Sage, London, 2007).
6. S. Kitayama, A. C. Snibbe, H. R. Markus, T. Suzuki, *Psychol. Sci.* **15**, 527 (2004).
7. Materials and methods are available as supporting material on Science Online.
8. We thank the Center for Advanced Study in the Behavioral Sciences and the R. C. Lee Charitable Foundation for support.

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/709/DC1

Materials and Methods
References and Notes

7 January 2010; accepted 6 April 2010
10.1126/science.1186799

Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA.

*To whom correspondence should be addressed. E-mail: spikelee@umich.edu

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/710>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/328/5979/710/DC1>

This article **cites 81 articles**, 29 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/328/5979/710#otherarticles>

This article has been **cited by** 1 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/cgi/content/full/328/5979/710#otherarticles>

This article appears in the following **subject collections**:

Immunology

<http://www.sciencemag.org/cgi/collection/immunology>

A Draft Sequence of the Neandertal Genome

Richard E. Green,^{1*}†† Johannes Krause,^{1†§} Adrian W. Briggs,^{1†§} Tomislav Maricic,^{1†§} Udo Stenzel,^{1†§} Martin Kircher,^{1†§} Nick Patterson,^{2†§} Heng Li,^{2†} Weiwei Zhai,^{3†||} Markus Hsi-Yang Fritz,^{4†} Nancy F. Hansen,^{5†} Eric Y. Durand,^{3†} Anna-Sapfo Malaspinas,^{3†} Jeffrey D. Jensen,^{6†} Tomas Marques-Bonet,^{7,13†} Can Alkan,^{7†} Kay Prüfer,^{1†} Matthias Meyer,^{1†} Hernán A. Burbano,^{1†} Jeffrey M. Good,^{1,8†} Rigo Schultz,¹ Ayinuer Aximu-Petri,¹ Anne Butthof,¹ Barbara Höber,¹ Barbara Höffner,¹ Madlen Siegemund,¹ Antje Weihmann,¹ Chad Nusbaum,² Eric S. Lander,² Carsten Russ,² Nathaniel Novod,² Jason Affourtit,⁹ Michael Egholm,⁹ Christine Verna,²¹ Pavao Rudan,¹⁰ Dejana Brajkovic,¹¹ Željko Kucan,¹⁰ Ivan Gušić,¹⁰ Vladimir B. Doronichev,¹² Liubov V. Golovanova,¹² Carlos Lalueza-Fox,¹³ Marco de la Rasilla,¹⁴ Javier Fortea,^{14†||} Antonio Rosas,¹⁵ Ralf W. Schmitz,^{16,17} Philip L. F. Johnson,^{18†} Evan E. Eichler,^{7†} Daniel Falush,^{19†} Ewan Birney,^{4†} James C. Mullikin,^{5†} Montgomery Slatkin,^{3†} Rasmus Nielsen,^{3†} Janet Kelso,^{1†} Michael Lachmann,^{1†} David Reich,^{2,20*}† Svante Pääbo^{1*†}

Neandertals, the closest evolutionary relatives of present-day humans, lived in large parts of Europe and western Asia before disappearing 30,000 years ago. We present a draft sequence of the Neandertal genome composed of more than 4 billion nucleotides from three individuals. Comparisons of the Neandertal genome to the genomes of five present-day humans from different parts of the world identify a number of genomic regions that may have been affected by positive selection in ancestral modern humans, including genes involved in metabolism and in cognitive and skeletal development. We show that Neandertals shared more genetic variants with present-day humans in Eurasia than with present-day humans in sub-Saharan Africa, suggesting that gene flow from Neandertals into the ancestors of non-Africans occurred before the divergence of Eurasian groups from each other.

The morphological features typical of Neandertals first appear in the European fossil record about 400,000 years ago (1–3). Progressively more distinctive Neandertal forms subsequently evolved until Neandertals disappeared from the fossil record about 30,000 years ago (4). During the later part of their history, Neandertals lived in Europe and Western Asia as far east as Southern Siberia (5) and as far south as the Middle East. During that time, Neandertals presumably came into contact with anatomically modern humans in the Middle East from at least 80,000 years ago (6, 7) and subsequently in Europe and Asia.

Neandertals are the sister group of all present-day humans. Thus, comparisons of the human genome to the genomes of Neandertals and apes allow features that set fully anatomically modern humans apart from other hominin forms to be identified. In particular, a Neandertal genome sequence provides a catalog of changes that have become fixed or have risen to high frequency in modern humans during the last few hundred thousand years and should be informative for identifying genes affected by positive selection since humans diverged from Neandertals.

Substantial controversy surrounds the question of whether Neandertals interbred with anatomically modern humans. Morphological features of present-day humans and early anatomically modern human fossils have been interpreted as evidence both for (8, 9) and against (10, 11) genetic exchange between Neandertals and the pre-

sumed ancestors of present-day Europeans. Similarly, analysis of DNA sequence data from present-day humans has been interpreted as evidence both for (12, 13) and against (14) a genetic contribution by Neandertals to present-day humans. The only part of the genome that has been examined from multiple Neandertals, the mitochondrial DNA (mtDNA) genome, consistently falls outside the variation found in present-day humans and thus provides no evidence for interbreeding (15–19). However, this observation does not preclude some amount of interbreeding (14, 19) or the possibility that Neandertals contributed other parts of their genomes to present-day humans (16). In contrast, the nuclear genome is composed of tens of thousands of recombining, and hence independently evolving, DNA segments that provide an opportunity to obtain a clearer picture of the relationship between Neandertals and present-day humans.

A challenge in detecting signals of gene flow between Neandertals and modern human ancestors is that the two groups share common ancestors within the last 500,000 years, which is no deeper than the nuclear DNA sequence variation within present-day humans. Thus, even if no gene flow occurred, in many segments of the genome, Neandertals are expected to be more closely related to some present-day humans than they are to each other (20). However, if Neandertals are, on average across many independent regions of the genome, more closely related to present-day humans in certain parts of the world than in others, this would strongly suggest that Neandertals ex-

changed parts of their genome with the ancestors of these groups.

Several features of DNA extracted from Late Pleistocene remains make its study challenging. The DNA is invariably degraded to a small average size of less than 200 base pairs (bp) (21, 22), it is chemically modified (21, 23–26), and extracts almost always contain only small amounts of endogenous DNA but large amounts of DNA from microbial organisms that colonized the specimens after death. Over the past 20 years, methods for ancient DNA retrieval have been developed (21, 22), largely based on the polymerase chain reaction (PCR) (27). In the case of the nuclear genome of Neandertals, four short gene sequences have been determined by PCR: fragments of the *MC1R* gene involved in skin pigmentation (28), a segment of the *FOXP2* gene involved in speech and language (29), parts of the ABO blood group locus (30), and a taste receptor gene (31). However, although PCR of ancient DNA can be multiplexed (32), it does not allow the retrieval of a large proportion of the genome of an organism.

The development of high-throughput DNA sequencing technologies (33, 34) allows large-scale, genome-wide sequencing of random pieces of DNA extracted from ancient specimens (35–37) and has recently made it feasible to sequence ge-

¹Department of Evolutionary Genetics, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany. ²Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA.

³Department of Integrative Biology, University of California, Berkeley, CA 94720, USA. ⁴European Molecular Biology Laboratory–European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK.

⁵Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA. ⁶Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, MA 01655, USA.

⁷Howard Hughes Medical Institute, Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA. ⁸Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA. ⁹454 Life Sciences, Branford, CT 06405, USA. ¹⁰Croatian Academy of Sciences and Arts, Zrinski trg 11, HR-10000 Zagreb, Croatia. ¹¹Croatian Academy of Sciences and Arts, Institute for Quaternary Paleontology and Geology, Ante Kovacic 5, HR-10000 Zagreb, Croatia. ¹²ANO Laboratory of Prehistory, St. Petersburg, Russia.

¹³Institute of Evolutionary Biology (UPF-CSIC), Dr. Aiguader 88, 08003 Barcelona, Spain. ¹⁴Area de Prehistoria Departamento de Historia Universidad de Oviedo, Oviedo, Spain. ¹⁵Departamento de Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁶Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁷Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁸Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹⁹Department of Microbiology, University College Cork, Cork, Ireland. ²⁰Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²¹Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²²Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²³Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²⁴Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²⁵Department of Biology, Emory University, Atlanta, GA 30322, USA. ²⁶Department of Microbiology, University College Cork, Cork, Ireland. ²⁷Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²⁸Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²⁹Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ³⁰Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ³¹Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ³²Department of Biology, Emory University, Atlanta, GA 30322, USA. ³³Department of Microbiology, University College Cork, Cork, Ireland. ³⁴Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ³⁵Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

³⁶Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ³⁷Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ³⁸Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ³⁹Department of Biology, Emory University, Atlanta, GA 30322, USA. ⁴⁰Department of Microbiology, University College Cork, Cork, Ireland. ⁴¹Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ⁴²Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

⁴³Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ⁴⁴Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ⁴⁵Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ⁴⁶Department of Biology, Emory University, Atlanta, GA 30322, USA. ⁴⁷Department of Microbiology, University College Cork, Cork, Ireland. ⁴⁸Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ⁴⁹Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

⁵⁰Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ⁵¹Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ⁵²Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ⁵³Department of Biology, Emory University, Atlanta, GA 30322, USA. ⁵⁴Department of Microbiology, University College Cork, Cork, Ireland. ⁵⁵Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ⁵⁶Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

⁵⁷Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ⁵⁸Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ⁵⁹Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ⁶⁰Department of Biology, Emory University, Atlanta, GA 30322, USA. ⁶¹Department of Microbiology, University College Cork, Cork, Ireland. ⁶²Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ⁶³Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

⁶⁴Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ⁶⁵Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ⁶⁶Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ⁶⁷Department of Biology, Emory University, Atlanta, GA 30322, USA. ⁶⁸Department of Microbiology, University College Cork, Cork, Ireland. ⁶⁹Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ⁷⁰Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

⁷¹Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ⁷²Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ⁷³Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ⁷⁴Department of Biology, Emory University, Atlanta, GA 30322, USA. ⁷⁵Department of Microbiology, University College Cork, Cork, Ireland. ⁷⁶Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ⁷⁷Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

⁷⁸Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ⁷⁹Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ⁸⁰Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ⁸¹Department of Biology, Emory University, Atlanta, GA 30322, USA. ⁸²Department of Microbiology, University College Cork, Cork, Ireland. ⁸³Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ⁸⁴Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

⁸⁵Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ⁸⁶Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ⁸⁷Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ⁸⁸Department of Biology, Emory University, Atlanta, GA 30322, USA. ⁸⁹Department of Microbiology, University College Cork, Cork, Ireland. ⁹⁰Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ⁹¹Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

⁹²Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ⁹³Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ⁹⁴Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ⁹⁵Department of Biology, Emory University, Atlanta, GA 30322, USA. ⁹⁶Department of Microbiology, University College Cork, Cork, Ireland. ⁹⁷Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ⁹⁸Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

⁹⁹Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁰⁰Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁰¹Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁰²Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹⁰³Department of Microbiology, University College Cork, Cork, Ireland. ¹⁰⁴Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁰⁵Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹⁰⁶Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁰⁷Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁰⁸Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁰⁹Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹¹⁰Department of Microbiology, University College Cork, Cork, Ireland. ¹¹¹Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹¹²Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹¹³Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹¹⁴Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹¹⁵Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹¹⁶Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹¹⁷Department of Microbiology, University College Cork, Cork, Ireland. ¹¹⁸Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹¹⁹Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹²⁰Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹²¹Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹²²Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹²³Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹²⁴Department of Microbiology, University College Cork, Cork, Ireland. ¹²⁵Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹²⁶Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹²⁷Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹²⁸Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹²⁹Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹³⁰Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹³¹Department of Microbiology, University College Cork, Cork, Ireland. ¹³²Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹³³Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹³⁴Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹³⁵Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹³⁶Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹³⁷Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹³⁸Department of Microbiology, University College Cork, Cork, Ireland. ¹³⁹Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁴⁰Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹⁴¹Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁴²Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁴³Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁴⁴Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹⁴⁵Department of Microbiology, University College Cork, Cork, Ireland. ¹⁴⁶Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁴⁷Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹⁴⁸Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁴⁹Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁵⁰Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁵¹Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹⁵²Department of Microbiology, University College Cork, Cork, Ireland. ¹⁵³Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁵⁴Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹⁵⁵Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁵⁶Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁵⁷Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁵⁸Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹⁵⁹Department of Microbiology, University College Cork, Cork, Ireland. ¹⁶⁰Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁶¹Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹⁶²Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁶³Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁶⁴Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁶⁵Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹⁶⁶Department of Microbiology, University College Cork, Cork, Ireland. ¹⁶⁷Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁶⁸Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹⁶⁹Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁷⁰Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁷¹Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁷²Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹⁷³Department of Microbiology, University College Cork, Cork, Ireland. ¹⁷⁴Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁷⁵Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹⁷⁶Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁷⁷Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁷⁸Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁷⁹Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹⁸⁰Department of Microbiology, University College Cork, Cork, Ireland. ¹⁸¹Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁸²Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹⁸³Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁸⁴Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁸⁵Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁸⁶Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹⁸⁷Department of Microbiology, University College Cork, Cork, Ireland. ¹⁸⁸Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁸⁹Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹⁹⁰Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁹¹Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁹²Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ¹⁹³Department of Biology, Emory University, Atlanta, GA 30322, USA. ¹⁹⁴Department of Microbiology, University College Cork, Cork, Ireland. ¹⁹⁵Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁹⁶Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

¹⁹⁷Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ¹⁹⁸Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ¹⁹⁹Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²⁰⁰Department of Biology, Emory University, Atlanta, GA 30322, USA. ²⁰¹Department of Microbiology, University College Cork, Cork, Ireland. ²⁰²Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²⁰³Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²⁰⁴Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²⁰⁵Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²⁰⁶Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²⁰⁷Department of Biology, Emory University, Atlanta, GA 30322, USA. ²⁰⁸Department of Microbiology, University College Cork, Cork, Ireland. ²⁰⁹Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²¹⁰Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²¹¹Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²¹²Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²¹³Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²¹⁴Department of Biology, Emory University, Atlanta, GA 30322, USA. ²¹⁵Department of Microbiology, University College Cork, Cork, Ireland. ²¹⁶Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²¹⁷Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²¹⁸Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²¹⁹Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²²⁰Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²²¹Department of Biology, Emory University, Atlanta, GA 30322, USA. ²²²Department of Microbiology, University College Cork, Cork, Ireland. ²²³Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²²⁴Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²²⁵Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²²⁶Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²²⁷Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²²⁸Department of Biology, Emory University, Atlanta, GA 30322, USA. ²²⁹Department of Microbiology, University College Cork, Cork, Ireland. ²³⁰Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²³¹Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²³²Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²³³Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²³⁴Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²³⁵Department of Biology, Emory University, Atlanta, GA 30322, USA. ²³⁶Department of Microbiology, University College Cork, Cork, Ireland. ²³⁷Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²³⁸Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²³⁹Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²⁴⁰Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²⁴¹Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²⁴²Department of Biology, Emory University, Atlanta, GA 30322, USA. ²⁴³Department of Microbiology, University College Cork, Cork, Ireland. ²⁴⁴Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²⁴⁵Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²⁴⁶Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²⁴⁷Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²⁴⁸Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²⁴⁹Department of Biology, Emory University, Atlanta, GA 30322, USA. ²⁵⁰Department of Microbiology, University College Cork, Cork, Ireland. ²⁵¹Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²⁵²Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²⁵³Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²⁵⁴Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²⁵⁵Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²⁵⁶Department of Biology, Emory University, Atlanta, GA 30322, USA. ²⁵⁷Department of Microbiology, University College Cork, Cork, Ireland. ²⁵⁸Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²⁵⁹Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²⁶⁰Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²⁶¹Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²⁶²Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²⁶³Department of Biology, Emory University, Atlanta, GA 30322, USA. ²⁶⁴Department of Microbiology, University College Cork, Cork, Ireland. ²⁶⁵Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²⁶⁶Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²⁶⁷Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²⁶⁸Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²⁶⁹Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²⁷⁰Department of Biology, Emory University, Atlanta, GA 30322, USA. ²⁷¹Department of Microbiology, University College Cork, Cork, Ireland. ²⁷²Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²⁷³Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²⁷⁴Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²⁷⁵Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²⁷⁶Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²⁷⁷Department of Biology, Emory University, Atlanta, GA 30322, USA. ²⁷⁸Department of Microbiology, University College Cork, Cork, Ireland. ²⁷⁹Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ²⁸⁰Department of Human Evolution, Max-Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany.

²⁸¹Department of Paleobiología, Museo Nacional de Ciencias Naturales, CSIC, Madrid, Spain. ²⁸²Der Landschaftsverband Rheinland–Landesmuseum Bonn, Bachstrasse 5-9, D-53115 Bonn, Germany. ²⁸³Abteilung für Vor- und Frühgeschichtliche Archäologie, Universität Bonn, Germany. ²⁸⁴Department of Biology, Emory University, Atlanta, GA 30322, USA. ²⁸⁵Department of Microbiology, University College Cork, Cork, Ireland. ²⁸⁶Department of Genetics, Harvard Medical School

nomes from late Pleistocene species (38). However, because a large proportion of the DNA present in most fossils is of microbial origin, comparison to genome sequences of closely related organisms is necessary to identify the DNA molecules that derive from the organism under study (39). In the case of Neandertals, the finished human genome sequence and the chimpanzee genome offer the opportunity to identify Neandertal DNA sequences (39, 40).

A special challenge in analyzing DNA sequences from the Neandertal nuclear genome is that most DNA fragments in a Neandertal are expected to be identical to present-day humans (41). Thus, contamination of the experiments with DNA from present-day humans may be mistaken for endogenous DNA. We first applied high-throughput sequencing to Neandertal specimens from Vindija Cave in Croatia (40, 42), a site from which cave bear remains yielded some of the first nuclear DNA sequences from the late Pleistocene in 1999 (43). Close to one million bp of nuclear DNA sequences from one bone were directly determined by high-throughput sequencing on the 454 platform (40), whereas DNA fragments from another extract from the same bone were cloned in a plasmid vector and used to sequence ~65,000 bp (42). These experiments, while demonstrating the feasibility of generating a Neandertal genome sequence, were preliminary in that they involved the transfer of DNA extracts prepared in a clean-room environment to conventional laboratories for processing and sequencing, creating an opportunity for contamination by present-day human DNA. Further analysis of the larger of these data sets (40) showed that it was contaminated with modern human DNA (44) to an extent of 11 to 40% (41). We employed a number of technical improvements, including the attachment of tagged sequence adaptors in the clean-room environment (23), to minimize the risk of contamination and determine about 4 billion bp from the Neandertal genome.

Paleontological samples. We analyzed a total of 21 Neandertal bones from Vindija Cave in Croatia that are of little morphological value. From below the surface of each of these bones, we removed 50 to 100 mg of bone powder using a sterile dentistry drill in our Leipzig clean-room facility. All samples were screened for the presence of Neandertal mtDNA by PCR, and three bones were selected for further analysis (Fig. 1A) [Supporting Online Material (SOM) Text 2]. The first of these bones, Vi33.16 (previously Vi-80) was discovered in stratigraphic layer G3 by Malez and co-workers in 1980 and has been directly dated by carbon-14 accelerator mass spectrometry to $38,310 \pm 2,130$ years before the present (B.P.) (uncalibrated) (19). It has been previously used for genome sequencing (40, 42) and for the determination of a complete mtDNA sequence (45). The second bone, Vi33.25, comes from layer I, which is deeper and thus older than layer G. A complete mtDNA sequence has been determined from this bone (15). It does not contain enough

collagen to allow a direct date. The third bone, Vi33.26, comes from layer G (sublayer unknown) and has not been previously used for large-scale DNA sequencing. It was directly dated to $44,450 \pm 550$ years B.P. (OxA-V-2291-18, uncalibrated).

Sequencing library construction. A total of nine DNA extracts were prepared from the three bones (table S4) using procedures to minimize laboratory contamination that we have developed over the past two decades (22, 41). Samples of each extract were used to construct Roche/454 sequencing libraries that carry the project-specific tag sequence 5'-TGAC-3' in their 3'-ends. Each library was amplified with the primers used in the 454 sequencing emulsion PCR process. To estimate the percentage of endogenous Neandertal DNA in the extracts, we carried out sequencing runs using the 454 Life Sciences GS FLX platform and mapped the reads against the human, chimpanzee, rhesus, and mouse genomes as well as all nucleotide sequences in GenBank. DNA sequences with a significantly better match to the primate genomes than to any of the other sources of sequences were further analyzed. Mitochondrial DNA contamination from modern humans was estimated by primer extension capture (46) using six biotinylated primers that target informative differences between human and Neandertal mtDNA (45), followed by sequencing on the GS FLX platform. Extracts that contained more than 1.5% hominin DNA relative to other DNA were used to construct further libraries. These were similarly analyzed to assess the percentage of hominin DNA and, if found suitable, were used for production sequencing on the 454 Life Sciences GS FLX/Titanium and Illumina GAII platforms.

Enrichment of Neandertal DNA. Depending on the extract, between 95 and 99% of the DNA sequenced in the libraries was derived from nonprimate organisms, which are presumably derived from microbes that colonized the bone after the death of the Neandertals. To improve the ratio of Neandertal to microbial DNA, we identified restriction enzymes that preferentially cut bacterial DNA sequences in the libraries and treated the libraries with these to increase the relative

proportion of Neandertal DNA in the libraries (SOM Text 1). Such enzymes, which have recognition sites rich in the dinucleotide CpG, allowed a 4- to 6-fold increase in the proportion of Neandertal DNA in the libraries sequenced. This is expected to bias the sequencing against GC-rich regions of the genome and is therefore not suitable for arriving at a complete Neandertal genome sequence. However, for producing an overview of the genome at about one-fold coverage, it drastically increases the efficiency of data production without unduly biasing coverage, especially in view of the fact that GC-rich sequences are over-represented in ancient DNA sequencing libraries (23, 45) so that the restriction enzyme treatment may help to counteract this bias.

Sequencing platforms and alignments. In the initial phase of the project, we optimized DNA extraction technology and library construction [e.g., (47)]. In a second phase, we carried out production sequencing on the 454 Life Sciences GS FLX platform from the bones Vi33.16 and Vi33.26 (0.5 Gb and 0.8 Gb of Neandertal sequence, respectively). In the third phase, we carried out production sequencing on the Illumina/Solexa GAII platform from the bones Vi33.16, Vi33.25, and Vi33.26 (1.2 Gb, 1.3 Gb, and 1.5 Gb, respectively) (table S4). Each molecule was sequenced from both ends (SOM Text 2), and bases were called with the machine learning algorithm Ibis (48). All reads were required to carry correct clean-room tags, and previous data where these tags were not used (40, 42) were not included in this study. Except when explicitly stated, the analyses below are based on the largest data sets, generated on the Illumina platform. In total, we generated 5.3 Gb of Neandertal DNA sequence from about 400 mg of bone powder. Thus, methods for extracting and sequencing DNA from ancient bones are now efficient enough to allow genome-wide DNA sequence coverage with relatively minor damage to well-preserved paleontological specimens.

The dominant type of nucleotide misincorporation when ancient DNA is amplified and sequenced is due to deamination of cytosine residues (25). This causes C to T transitions in the DNA sequences,

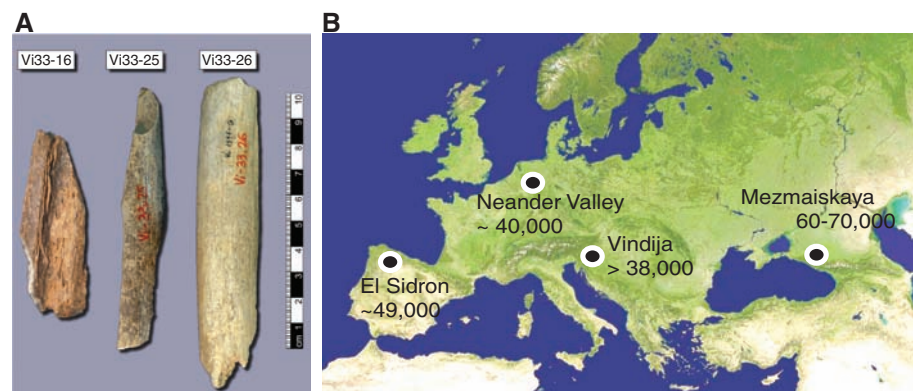


Fig. 1. Samples and sites from which DNA was retrieved. (A) The three bones from Vindija from which Neandertal DNA was sequenced. (B) Map showing the four archaeological sites from which bones were used and their approximate dates (years B.P.).

particularly toward the 5'-ends of DNA reads, where at the first position ~40% of cytosine residues can appear as thymine residues. The frequency of C to T misincorporations progressively diminishes further into the molecules. At the 3'-ends, complementary G to A transitions are seen as a result of the enzymatic fill-in procedure in which blunt DNA ends are created before adaptor ligation (23). We implemented an alignment approach that takes these nucleotide misincorporation patterns into account (SOM Text 3) and aligned the Neandertal sequences to either the reference human genome (UCSC hg18), the reference chimpanzee genome (*panTro2*), or the inferred human-chimpanzee common ancestral sequence (SOM Text 3).

To estimate the error rate in the Neandertal DNA sequences determined, we compared reads that map to the mitochondrial genomes, which we assembled to 35-, 29- and 72-fold coverage for each of the bones, respectively (15, 45) (SOM Text 4). Although C to T and G to A substitutions, which are caused by deaminated cytosine residues, occur at a rate of 4.5 to 5.9%, other error rates are at most 0.3% (fig. S4). Because we sequence each DNA fragment from both sides, and most fragments more than once (49), the latter error rate is substantially lower than the error rate of the Illumina platform itself (48, 50).

Number of Neandertal individuals. To assess whether the three bones come from different individuals, we first used their mtDNAs. We have previously determined the complete mtDNA sequences from the bones Vi33.16 and Vi33.25 (15, 45), and these differ at 10 positions. Therefore, Vi33.16 and Vi33.25 come from different Neandertal individuals. For the bone Vi33.26, we assembled the mtDNA sequence (SOM Text 4) and found it to be indistinguishable from Vi33.16, suggesting that it could come from the same individual. We analyzed autosomal DNA sequences from the three bones (SOM Text 4) by asking whether the frequency of nucleotide differences between pairs of bones was significantly higher than the frequency of differences within the bones. We find that the within-bone differences are significantly fewer than the between-bone differences for all three comparisons ($P \leq 0.001$ in all cases). Thus, all three bones derive from different individuals, although Vi33.16 and Vi33.26 may stem from maternally related individuals.

Estimates of human DNA contamination.

We used three approaches that target mtDNA, Y chromosomal DNA, and nuclear DNA, respectively, to gauge the ratio of present-day human relative to Neandertal DNA in the data produced. To analyze the extent of mtDNA contamination, we used the complete mtDNA from each bone to identify positions differing from at least 99% of a worldwide panel of 311 contemporary human mtDNAs, ignoring positions where a substitution in the sequences from the Neandertal library could be due to cytosine deamination (45). For each sequencing library, the DNA fragments that cover these positions were then classified according to whether they appear to be of Neandertal or modern human origin (SOM Text 5 and table S15). For each bone, the level of mtDNA contamination is estimated to be below 0.5% (Table 1).

Because prior to this study no fixed differences between Neandertal and present-day humans in the nuclear genome were known, we used two alternative strategies to estimate levels of nuclear contamination. In the first strategy, we determined the sex of the bones. For bones derived from female Neandertals, we then estimated modern human male DNA contamination by looking for the presence of Y chromosomal DNA fragments (SOM Text 6). For this purpose, we identified 111,132 nucleotides in the nonrecombining parts of the human reference Y chromosome that are located in contiguous DNA segments of at least 500 nucleotides, carry no repetitive elements, and contain no 30-nucleotide oligomer elsewhere in the genome with fewer than three mismatches. Between 482 and 611 such fragments would be expected for a male Neandertal bone. However, only 0 to 4 fragments are observed (Table 1). We conclude that the three bones are all from female Neandertals and that previous suggestions that Vi33.16 was a male (40, 42) were due to mismatching of autosomal and X chromosomal reads to the Y chromosome. We estimate the extent of DNA contamination from modern human males in the combined data to be about 0.60%, with an upper 95% bound of 1.53%.

In the second strategy, we take advantage of the fact that sites where present-day humans carry a high frequency of a derived allele (i.e., not seen in chimpanzee) while Neandertals carry a high frequency of the ancestral allele (i.e., matching the

chimpanzee) provide information about the extent of contamination. To implement this idea, we identified sites where five present-day humans that we sequenced (see below) all differ from the chimpanzee genome by a transversion. We further restricted the analysis to sites covered by two fragments in one Neandertal and one fragment in another Neandertal and where at least one ancestral allele was seen in both individuals. The additional fragment from the first Neandertal then provides an estimate of contamination in combination with heterozygosity at this class of sites (Table 1). Using these data (SOM Text 7), we derive a maximum likelihood estimate of contamination of 0.7% with an upper 95% bound of 0.8%.

In summary, all three measurements of human mtDNA contamination produce estimates of less than 1% contamination. Thus, the vast majority of these data represent bona fide Neandertal DNA sequences.

Average DNA divergence between Neandertals

and humans. To estimate the DNA sequence divergence per base pair between the genomes of Neandertals and the reference human genome sequence, we generated three-way alignments between the Neandertal, human, and chimpanzee genomes, filtering out genomic regions that may be duplicated in either humans or chimpanzees (SOM Text 10) and using an inferred genome sequence of the common ancestor of humans and chimpanzees as a reference (51) to avoid potential biases (39). We then counted the number of substitutions specific to the Neandertal, the human, and the chimpanzee genomes (Fig. 2). The overall number of substitutions unique to the Neandertal genome is about 30 times as high as on the human lineage. Because these are largely due to transitions resulting from deamination of cytosine residues in the Neandertal DNA, we restricted the divergence estimates to transversions. We then observed four to six times as many on the Neandertal as on the human lineage, probably due to sequencing errors in the low-coverage Neandertal DNA sequences. The numbers of transversions on the human lineage, as well as those on the lineage from the Neandertal-human ancestor to the chimpanzee, were used to estimate the average divergence between DNA sequences in Neandertals and present-day humans, as a fraction of the lineage from the human reference genome to the common

Table 1. Estimates of human DNA contamination in the DNA sequences produced. Numbers in bold indicate summary contamination estimates over all Vindija data.

	mtDNA contamination				Y chromosomal contamination				Neandertal diversity (1/2) plus contamination*			Nuclear ML contamination
	Human	Neandertal	Percent	95% C.I.	Observed	Expected	Percent	95% C.I.	Percent	Upper	95% C.I.	Percent (95% C.I.)
Vi33.16	56	20,456	0.27	0.21–0.35	4	255	1.57	0.43–3.97	1.4	2.2		n/a
Vi33.25	7	1,691	0.41	0.17–0.85	0	201	0.0	0.00–1.82	1.0	1.7		n/a
Vi33.26	10	4,810	0.21	0.10–0.38	0	210	0.0	0.00–1.74	1.1	1.9		n/a
All data	73	26,957	0.27	0.21–0.34	4	666	0.60	0.16–1.53	1.2	1.6		0.7 (0.6–0.8)

*Assuming similar extents of contamination in the three bones and that individual heterozygosity and population nucleotide diversity is the same for this class of sites.

ancestor of Neandertals, humans, and chimpanzees. For autosomes, this was 12.7% for each of the three bones analyzed. For the X chromosome, it was 11.9 to 12.4% (table S26). Assuming an average DNA divergence of 6.5 million years between the human and chimpanzee genomes (52), this results in a point estimate for the average divergence of Neandertal and modern human autosomal DNA sequences of 825,000 years. We caution that this is only a rough estimate because of the uncertainty about the time of divergence of humans and chimpanzees.

Additional Neandertal individuals. To put the divergence of the Neandertal genome sequences from Vindija Cave into perspective with regard to other Neandertals, we generated a much smaller amount of DNA sequence data from three Neandertal bones from three additional sites (SOM Text 8) that cover much of the geographical range of late Neandertals (Fig. 1B): El Sidron in Asturias, Spain, dated to ~49,000 years B.P. (53); Feldhofer Cave in the Neander Valley, Germany, from which we sequenced the type specimen found in 1856 dated to ~42,000 years B.P. (54); and Mezmaiskaya Cave in the Caucasus, Russia, dated to 60,000 to 70,000 years B.P. (55). DNA divergences estimated for each of these specimens to the human reference genome (table S26) show that none of them differ significantly from the Vindija individuals, although these estimates are relatively uncertain due to the limited amount of DNA sequence

data. It is noteworthy that the Mezmaiskaya specimen, which is 20,000 to 30,000 years older than the other Neandertals analyzed and comes from the easternmost location, does not differ in divergence from the other individuals. Thus, within the resolution of our current data, Neandertals from across a great part of their range in western Eurasia are equally related to present-day humans.

Five present-day human genomes. To put the divergence of the Neandertal genomes into perspective with regard to present-day humans, we sequenced the genomes of one San from Southern Africa, one Yoruba from West Africa, one Papua New Guinean, one Han Chinese, and one French from Western Europe to 4- to 6-fold coverage on the Illumina GAI platform (SOM Text 9). These sequences were aligned to the chimpanzee and human reference genomes and analyzed using a similar approach to that used for the Neandertal data. Autosomal DNA sequences of these individuals diverged 8.2 to 10.3% back along the lineage leading to the human reference genome, considerably less than the 12.7% seen in Neandertals (SOM Text 10). We note that the divergence estimate for the Yoruba individual to the human genome sequence is ~14% greater than previous estimates for an African American individual (56) and similarly greater than the heterozygosity measured in another Yoruba individual (33). This may be due to differences in the alignment and filtering procedures between

this and previous studies (SOM Text 9 and 10). Nevertheless, the divergence of the Neandertal genome to the human reference genome is greater than for any of the present-day human genomes analyzed.

Distributions of DNA divergences to humans.

To explore the variation of DNA sequence divergence across the genome, we analyzed the divergence of the Neandertals and the five humans to the reference human genome in 100 kilobase windows for which at least 50 informative transversions were observed. The majority of the Neandertal divergences overlap with those of the humans (Fig. 3), reflecting the fact that Neandertals fall inside the variation of present-day humans. However, the overall divergence is greater for the three Neandertal genomes. For example, their modes are around divergences of ~11%, whereas for the San the mode is ~9% and for the other present-day humans ~8%. For the Neandertals, 13% of windows have a divergence above 20%, whereas this is the case for 2.5% to 3.7% of windows in the current humans.

Furthermore, whereas in the French, Han, and Papuan individuals, 9.8%, 7.8%, and 5.9% of windows, respectively, show between 0% and 2% divergence to the human reference genome, in the San and the Yoruba this is the case for 1.7% and 3.7%, respectively. For the three Neandertals, 2.2 to 2.5% of windows show 0% to 2% divergence to the reference genome.

A catalog of features unique to the human genome. The Neandertal genome sequences allow us to identify features unique to present-day humans relative to other, now extinct, hominins. Of special interest are features that may have functional consequences. We thus identified, from whole genome alignments, sites where the human genome reference sequence does not match chimpanzee, orangutan, and rhesus macaque. These are likely to have changed on the human lineage since the common ancestor with chimpanzee. Where Neandertal fragments overlapped, we constructed consensus sequences and joined them into "minicontigs," which were used to determine the Neandertal state at the positions that changed

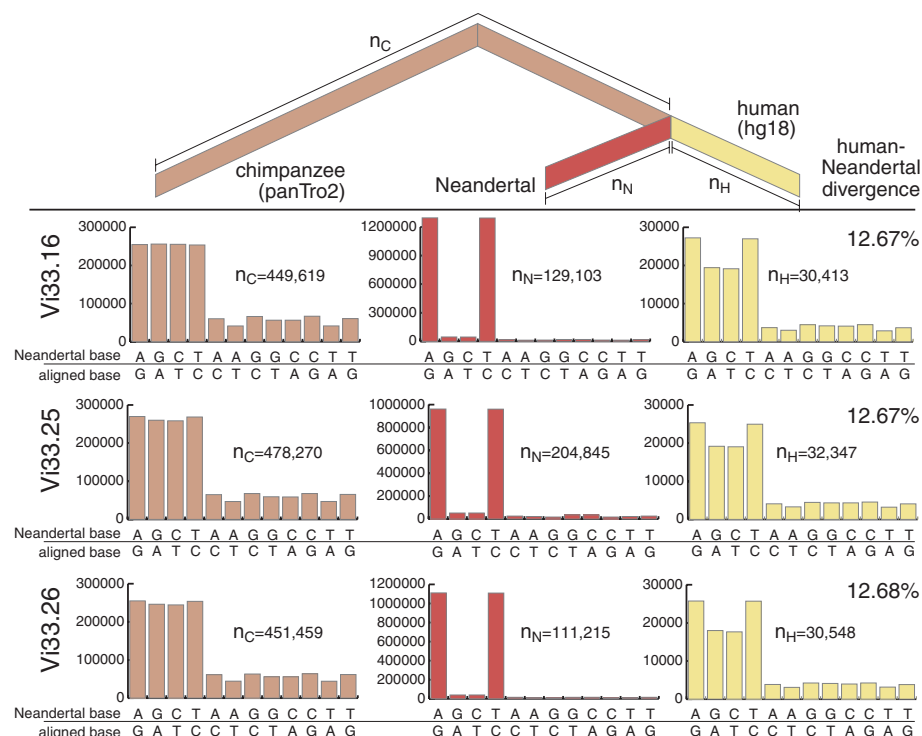


Fig. 2. Nucleotide substitutions inferred to have occurred on the evolutionary lineages leading to the Neandertals, the human, and the chimpanzee genomes. In red are substitutions on the Neandertal lineage, in yellow the human lineage, and in pink the combined lineage from the common ancestor of these to the chimpanzee. For each lineage and each bone from Vindija, the distributions and numbers of substitutions are shown. The excess of C to T and G to A substitutions are due to deamination of cytosine residues in the Neandertal DNA.

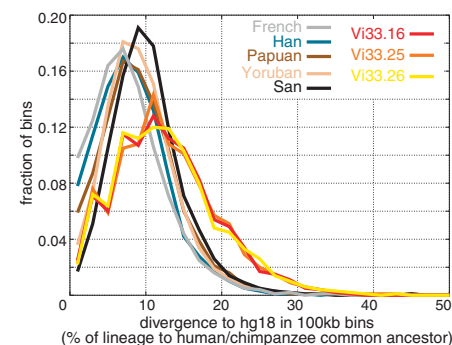


Fig. 3. Divergence of Neandertal and human genomes. Distributions of divergence from the human genome reference sequence among segments of 100 kb are shown for three Neandertals and the five present-day humans.

Table 2. Amino acid changes that are fixed in present-day humans but ancestral in Neandertals. The table is sorted by Grantham scores (GS). Based on the classification proposed by Li *et al.* in (87), 5 amino acid substitutions are radical (>150), 7 moderately radical (101 to 150), 33 moderately conservative (51 to

100) and 32 conservative (1 to 50). One substitution creates a stop codon. Genes showing multiple substitutions have bold SwissProt identifiers. (Table S15 shows the human and chimpanzee genome coordinates, additional database identifiers, and the respective bases.) Genes with two fixed amino acids are indicated in bold.

ID	Pos	AA	GS	Description/function
RPTN	785	*R	–	Multifunctional epidermal matrix protein
GREB1	1164	R/C	180	Response gene in estrogen receptor–regulated pathway
OR1K1	267	R/C	180	Olfactory receptor, family 1, subfamily K, member 1
SPAG17	431	Y/D	160	Involved in structural integrity of sperm central apparatus axoneme
NLRX1	330	Y/D	160	Modulator of innate immune response
NSUN3	78	S/F	155	Protein with potential SAM-dependent methyl-transferase activity
RGS16	197	D/A	126	Retinally abundant regulator of G-protein signaling
BOD1L	2684	G/R	125	Biorientation of chromosomes in cell division 1-like
CF170	505	S/C	112	<i>Uncharacterized protein: C6orf170</i>
STEAP1	336	C/S	112	Metalloreductase, six transmembrane epithelial antigen of prostate 1
F16A2	630	R/S	110	<i>Uncharacterized protein: family with sequence similarity 160, member A2</i>
LYN	569	R/S	110	Leukocyte receptor tyrosine kinase
BEND2	261	V/G	109	<i>Uncharacterized protein: BEN domain-containing protein 2</i>
O52W1	51	P/L	98	Olfactory receptor, family 52, subfamily W, member 1
CAN15	427	L/P	98	Small optic lobes homolog, linked to visual system development
SCAP	140	I/T	89	Escort protein required for cholesterol as well as lipid homeostasis
TTF1	474	I/T	89	RNA polymerase I termination factor
OR5K4	175	H/D	81	Olfactory receptor, family 5, subfamily K, member 4
SCML1	202	T/M	81	Putative polycomb group (PcG) protein
TTL10	394	K/T	78	Probable tubulin polyglutamylase, forming polyglutamate side chains on tubulin
AFF3	516	S/P	74	Putative transcription activator, function in lymphoid development/oncogenesis
EYA2	131	S/P	74	Tyrosine phosphatase, dephosphorylating “Tyr-142” of histone H2AX
NOP14	493	T/R	71	Involved in nucleolar processing of pre-18S ribosomal RNA
PRDM10	1129	N/T	65	PR domain containing 10, may be involved in transcriptional regulation
BTLA	197	N/T	65	B and T lymphocyte attenuator
O2AT4	224	V/A	64	Olfactory receptor, family 2, subfamily AT, member 4
CAN15	356	V/A	64	Small optic lobes homolog, linked to visual system development
ACCN4	160	V/A	64	Amiloride-sensitive cation channel 4, expressed in pituitary gland
PUR8	429	V/A	64	Adenylsuccinate lyase (purine synthesis)
MCHR2	324	A/V	64	Receptor for melanin-concentrating hormone, coupled to G proteins
AHR	381	V/A	64	Aromatic hydrocarbon receptor, a ligand-activated transcriptional activator
FAAH1	476	A/G	60	Fatty acid amide hydrolase
SPAG17	1415	T/A	58	Involved in structural integrity of sperm central apparatus axoneme
ZF106	697	A/T	58	Zinc finger protein 106 homolog / SH3-domain binding protein 3
CAD16	342	T/A	58	Calcium-dependent, membrane-associated glycoprotein (cellular recognition)
K1C16	306	T/A	58	Keratin, type I cytoskeletal 16 (expressed in esophagus, tongue, hair follicles)
LIMS2	360	T/A	58	Focal adhesion protein, modulates cell spreading and migration
ZN502	184	T/A	58	Zinc finger protein 502, may be involved in transcriptional regulation
MEPE	391	A/T	58	Matrix extracellular phosphoglycoprotein, putative role in mineralization
FSTL4	791	T/A	58	Follistatin-related protein 4 precursor
SNTG1	241	T/S	58	Syntrophin, gamma 1; binding/organizing subcellular localization of proteins
RPTN	735	K/E	56	Multifunctional epidermal matrix protein
BCL9L	543	S/G	56	Nuclear cofactor of beta-catenin signaling, role in tumorigenesis
SSH2	1033	S/G	56	Protein phosphatase regulating actin filament dynamics
PEG3	1521	S/G	56	Apoptosis induction in cooperation with SIAH1A
DJC28	290	K/Q	53	DnaJ (Hsp40) homolog, may have role in protein folding or as a chaperone
CLTR2	50	F/V	50	Receptor for cysteinyl leukotrienes, role in endocrine and cardiovascular systems
KIF15	827	N/S	46	Putative kinesin-like motor enzyme involved in mitotic spindle assembly
SPOC1	355	Q/R	43	<i>Uncharacterized protein: SPOC domain containing 1</i>
TTF1	229	R/Q	43	RNA polymerase I termination factor
F166A	134	T/P	38	<i>Uncharacterized protein: family with sequence similarity 166, member A</i>
CL066	426	V/L	32	<i>Uncharacterized protein: chromosome 12 open reading frame 66</i>
PCD16	763	E/Q	29	Calcium-dependent cell-adhesion protein, fibroblasts expression
TRPM5	1088	I/V	29	Voltage-modulated cation channel (VCAM), central role in taste transduction
S36A4	330	H/R	29	Solute carrier family 36 (proton/amino acid symporter)
GP132	328	E/Q	29	High-affinity G-protein couple receptor for lysophosphatidylcholine (LPC)
ZFY26	237	H/R	29	Zinc finger FYVE domain-containing, associated with spastic paraplegia-15

continued on next page

ID	Pos	AA	GS	Description/function
CALD1	671	I/V	29	Actin- and myosin-binding protein, regulation of smooth muscle contraction
CDC2A	606	I/V	29	Regulator of chromosome structure during mitosis
GPAA1	275	E/Q	29	Glycosylphosphatidylinositol anchor attachment protein
ARSF	200	I/V	29	Arylsulfatase F precursor, relevant for composition of bone and cartilage matrix
OR4D9	303	R/K	26	Olfactory receptor, family 4, subfamily D, member 9
EMIL2	155	R/K	26	Elastin microfibril interface-located protein (smooth muscle anchoring)
PHLP	216	K/R	26	Putative modulator of heterotrimeric G proteins
TKTL1	317	R/K	26	Transketolase-related protein
MIIP	280	H/Q	24	Inhibits glioma cells invasion, down-regulates adhesion and motility genes
SPTA1	265	N/D	23	Constituent of cytoskeletal network of the erythrocyte plasma membrane
PCD16	777	D/N	23	Calcium-dependent cell-adhesion protein, fibroblasts expression
CS028	326	L/F	22	<i>Uncharacterized protein</i> : chromosome 19 open reading frame 28
PIGZ	425	L/F	22	Mannosyltransferase for glycosylphosphatidylinositol-anchor biosynthesis
DISP1	1079	V/M	21	Segment-polarity gene required for normal Hedgehog (Hh) signaling
RNA57	44	M/V	21	Protein with RNase activity for broad-spectrum of pathogenic microorganisms
KR241	205	V/M	21	Keratin-associated protein, formation of a rigid and resistant hair shaft
SPLC3	108	I/M	10	Short palate, lung, and nasal epithelium carcinoma-associated protein
NCOA6	823	I/M	10	Hormone-dependent coactivation of several receptors
WWC2	479	M/I	10	<i>Uncharacterized protein</i> : WW, C2, and coiled-coil domain containing 2
ASCC1	301	E/D	0	Enhancer of NF-kappa-B, SRF, and AP1 transactivation
PROM2	458	D/E	0	Plasma membrane protrusion in epithelial and nonepithelial cells

on the human lineage. To minimize alignment errors and substitutions, we disregarded all substitutions and insertions or deletions (indels) within 5 nucleotides of the ends of minicontigs or within 5 nucleotides of indels.

Among 10,535,445 substitutions and 479,863 indels inferred to have occurred on the human lineage, we have information in the Neandertal genome for 3,202,190 and 69,029, i.e., 30% and 14%, respectively. The final catalog thus represents those sequenced positions where we have high confidence in their Neandertal state (SOM Text 11). As expected, the vast majority of those substitutions and indels (87.9% and 87.3%, respectively) occurred before the Neandertal divergence from modern humans.

Features that occur in all present-day humans (i.e., have been fixed), although they were absent or variable in Neandertals, are of special interest. We found 78 nucleotide substitutions that change the protein-coding capacity of genes where modern humans are fixed for a derived state and where Neandertals carry the ancestral (chimpanzee-like) state (Table 2 and table S28). Thus, relatively few amino acid changes have become fixed in the last few hundred thousand years of human evolution; an observation consistent with a complementary study (57). We found only five genes with more than one fixed substitution changing the primary structure of the encoded proteins. One of these is *SPAG17*, which encodes a protein important for the axoneme, a structure responsible for the beating of the sperm flagellum (58). The second is *PCD16*, which encodes fibroblast cadherin-1, a calcium-dependent cell-cell adhesion molecule that may be involved in wound healing (59). The third is *TTF1*, a transcription termination factor that regulates ribosomal gene transcription (60). The fourth is *CAN15*, which encodes a protein of unknown function. The fifth is *RPTN*, which encodes repetin,

an extracellular epidermal matrix protein (61) that is expressed in the epidermis and at high levels in eccrine sweat glands, the inner sheaths of hair roots, and the filiform papilli of the tongue.

One of the substitutions in *RPTN* creates a stop codon that causes the human protein to contain 784 rather than 892 amino acids (SOM Text 11). We identified no fixed start codon differences, although the start codon in the gene *TRPM1* that is present in Neandertals and chimpanzees has been lost in some present-day humans. *TRPM1* encodes melastatin, an ion channel important for maintaining melanocyte pigmentation in the skin. It is intriguing that skin-expressed genes comprise three out of six genes that either carry multiple fixed substitutions changing amino acids or in which a start or stop codon has been lost or gained. This suggests that selection on skin morphology and physiology may have changed on the hominin lineage.

We also identified a number of potential regulatory substitutions that are fixed in present-day humans but not Neandertals. Specifically, we find 42 substitutions and three indels in 5'-untranslated regions, and 190 substitutions and 33 indels in 3'-untranslated regions that have become fixed in humans since they diverged from Neandertals. Of special interest are microRNAs (miRNAs), small RNAs that regulate gene expression by mRNA cleavage or repression of translation. We found one miRNA where humans carry a fixed substitution at a position that was ancestral in Neandertals (*hsa-mir-1304*) and one case of a fixed single nucleotide insertion where Neandertal is ancestral (*AC109351.3*). While the latter insertion is in a bulge in the inferred secondary structure of the miRNA that is unlikely to affect folding or putative targets, the substitution in *mir-1304* occurs in the seed region, suggesting that it is likely to have altered target specificity in modern humans relative to Neandertals and other apes (fig. S16).

Human accelerated regions (HARs) are defined as regions of the genome that are conserved throughout vertebrate evolution but that changed radically since humans and chimpanzees split from their common ancestor. We examined 2613 HARs (SOM Text 11) and obtained reliable Neandertal sequence for 3259 human-specific changes in HARs. The Neandertals carry the derived state at 91.4% of these, significantly more than for other human-specific substitutions and indels (87.9%). Thus, changes in the HARs tend to predate the split between Neandertals and modern humans. However, we also identified 51 positions in 45 HARs where Neandertals carry the ancestral version whereas all known present-day humans carry the derived version. These represent recent changes that may be particularly interesting to explore functionally.

Neandertal segmental duplications. We analyzed Neandertal segmental duplications by measuring excess read-depth to identify and predict the copy number of duplicated sequences, defined as those with >95% sequence identity (62). A total of 94 Mb of segmental duplications were predicted in the Neandertal genome (table S33), which is in close agreement with what has been found in present-day humans (62) (fig. S18). We identified 111 potentially Neandertal-specific segmental duplications (average size 22,321 bp and total length 1862 kb) that did not overlap with human segmental duplications (fig. S20). Although direct experimental validation is not possible, we note that 81% (90/111) of these regions also showed excess sequence diversity (>3 SD beyond the mean) consistent with their being bona fide duplications (fig. S21). Many of these regions also show some evidence of increased copy number in humans, although they have not been previously classified as duplications (fig. S22). We identified only three putative Neandertal-specific

duplications with no evidence of duplication among humans or any other primate (fig. S23), and none contained known genes.

A comparison to any single present-day human genome reveals that 89% of the detected duplications are shared with Neandertals. This is lower than the proportion seen between present-day humans (around 95%) but higher than what is observed when the Neandertals are compared with the chimpanzee (67%) (fig. S19).

Because the Neandertal data set is derived from a pool of three individuals and represents an average sequence coverage of 1.3-fold after filtering, we created two resampled sets from three human genomes (SOM Text 12) at a comparable level of mixture and coverage (table S34 and figs. S24 and S25). The analysis of both resampled sets show a nonsignificant trend toward more duplicated sequences among Neandertals than among present-day humans (88,869 kb, $N = 1129$ regions for present-day humans versus 94,419 kb, $N = 1194$ for the Neandertals) (fig. S25).

We also estimated the copy number for Neandertal genes and compared it with those from

three previously analyzed human genomes (SOM Text 12). Copy number was correlated between the two groups ($r^2 = 0.91$) (fig. S29), with only 43 genes (15 nonredundant genes >10 kb) showing a difference of more than five copies (tables S35 and S36). Of these genes, 67% (29/43) are increased in Neandertals compared with present-day humans, and most of these are genes of unknown function. One of the most extreme examples is the gene *PRR20* (NM_198441), for which we predicted 68 copies in Neandertals, 16 in humans, and 58 in the chimpanzee. It encodes a hypothetical proline-rich protein of unknown function. Other genes with predicted higher copy number in humans as opposed to Neandertals included *NBPFL4* (*DUF1220*), *DUX4* (NM_172239), *REXO1L1* (NM_033178), and *TBC1D3* (NM_001123391).

A screen for positive selection in early modern humans. Neandertals fall within the variation of present-day humans for many regions of the genome; that is, Neandertals often share derived single-nucleotide polymorphism (SNP) alleles with present-day humans. We devised an approach to detect positive selection in early modern humans

that takes advantage of this fact by looking for genomic regions where present-day humans share a common ancestor subsequent to their divergence from Neandertals, and Neandertals therefore lack derived alleles found in present-day humans (except in rare cases of parallel substitutions) (Fig. 4A). Gene flow between Neandertals and modern humans after their initial population separation might obscure some cases of positive selection by causing Neandertals and present-day humans to share derived alleles, but it will not cause false-positive signals.

We identified SNPs as positions that vary among the five present-day human genomes of diverse ancestry plus the human reference genome and used the chimpanzee genome to determine the ancestral state (SOM Text 13). We ignored SNPs at CpG sites since these evolve rapidly and may thus be affected by parallel mutations. We identified 5,615,438 such SNPs, at about 10% of which Neandertals carry the derived allele. As expected, SNPs with higher frequencies of the derived allele in present-day humans were more likely to show the derived allele in Neandertals

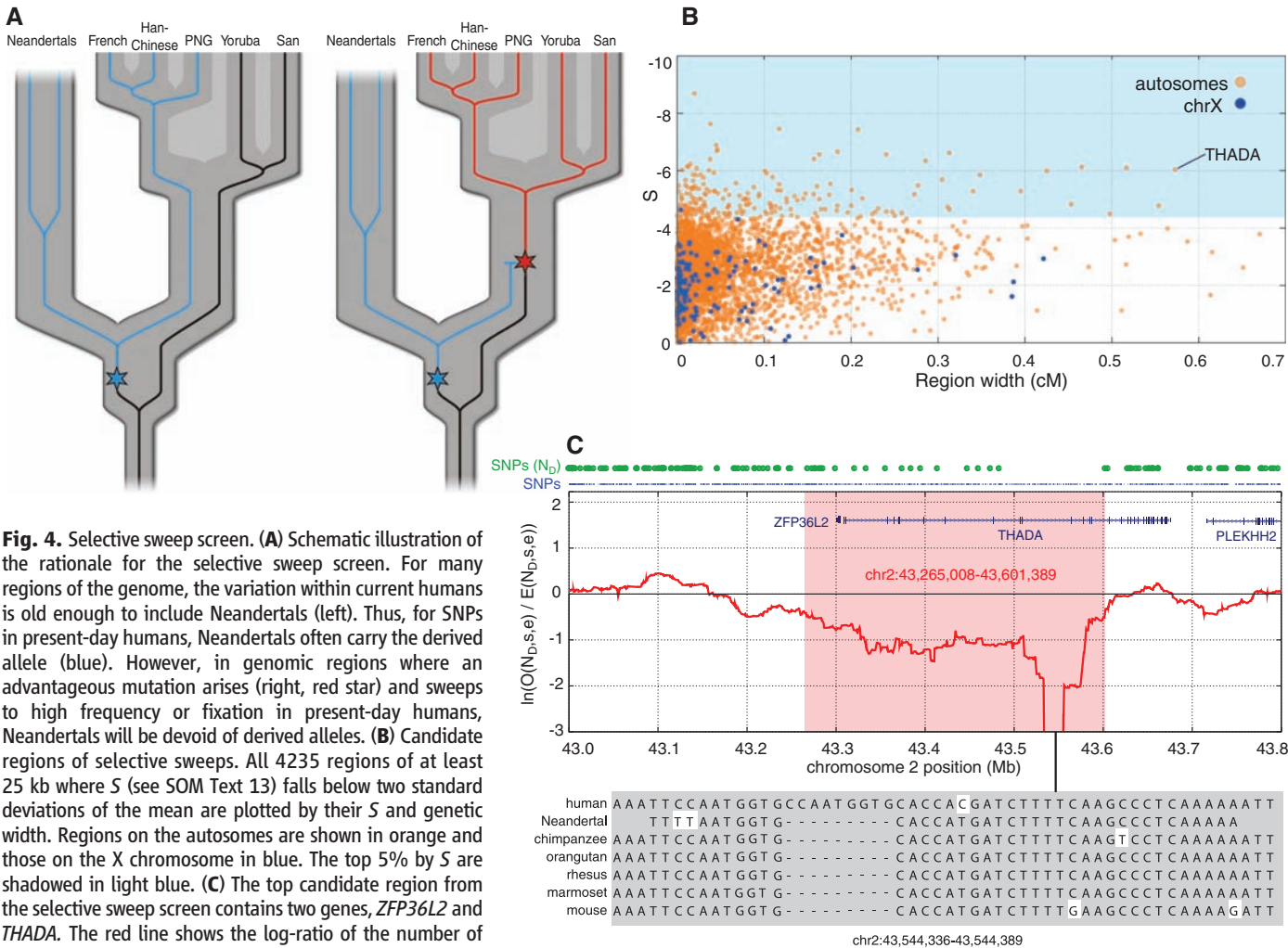


Fig. 4. Selective sweep screen. **(A)** Schematic illustration of the rationale for the selective sweep screen. For many regions of the genome, the variation within current humans is old enough to include Neandertals (left). Thus, for SNPs in present-day humans, Neandertals often carry the derived allele (blue). However, in genomic regions where an advantageous mutation arises (right, red star) and sweeps to high frequency or fixation in present-day humans, Neandertals will be devoid of derived alleles. **(B)** Candidate regions of selective sweeps. All 4235 regions of at least 25 kb where S (see SOM Text 13) falls below two standard deviations of the mean are plotted by their S and genetic width. Regions on the autosomes are shown in orange and those on the X chromosome in blue. The top 5% by S are shadowed in light blue. **(C)** The top candidate region from the selective sweep screen contains two genes, *ZFP36L2* and *THADA*. The red line shows the log-ratio of the number of observed Neandertal-derived alleles versus the number of expected Neandertal-derived alleles, within a 100 kilobase window. The blue dots above the panel indicate all SNP positions, and the green dots indicate SNPs where the Neandertal carries the derived allele.

(fig. S31A). We took advantage of this fact to calculate (fig. S31C) the expected number of Neandertal-derived alleles within a given region of the human genome. The observed numbers of derived alleles were then compared with the expected numbers to identify regions where the Neandertal carries fewer derived alleles than expected relative to the human allelic states. A unique feature of this method is that it has more power to detect older selective sweeps where allele frequency spectra in present-day humans have recovered to the point that appreciable derived allele frequencies are observed, whereas it has relatively low power to detect recent selective sweeps where the derived alleles are at low frequencies in present-day humans. It is therefore particularly suited to detect positive selection that occurred early during the history of modern human ancestors in conjunction with, or shortly after, their population divergence from Neandertals (Fig. 4A).

We identified a total of 212 regions containing putative selective sweeps (Fig. 4B and SOM Text S37). The region with the strongest statistical signal contained a stretch of 293 consecutive SNP positions in the first half of the gene *AUTS2* where only ancestral alleles are observed in the Neandertals (fig. S34).

We ranked the 212 regions with respect to their genetic width in centimorgans (Fig. 4B, and table S37) because the size of a region affected by a selective sweep will be larger the fewer generations it took for the sweep to reach fixation, as fewer recombination events will then have occurred during the sweep. Thus, the more intense the selection that drove a putative sweep, the larger the affected region is expected to be. Table 3 lists the 20 widest regions and the genes encoded in them. Five of the regions contain no protein-coding

genes. These may thus contain structural or regulatory genomic features under positive selection during early human history. The remaining 15 regions contain between one and 12 genes. The widest region is located on chromosome 2 and contains the gene *THADA*, where a region of 336 kb is depleted of derived alleles in Neandertals. SNPs in the vicinity of *THADA* have been associated with type II diabetes, and *THADA* expression differs between individuals with diabetes and healthy controls (63). Changes in *THADA* may thus have affected aspects of energy metabolism in early modern humans. The largest deficit of derived alleles in Neandertal *THADA* is in a region where the Neandertals carry ancestral alleles at 186 consecutive human SNP positions (Fig. 4C). In this region, we identified a DNA sequence element of ~700 bp that is conserved from mouse to primates, whereas the human reference genome as well as the four humans for which data are available carry an insertion of 9 bp that is not seen in the Neandertals. We note, however, that this insertion is polymorphic in humans, as it is in dbSNP.

Mutations in several genes in Table 3 have been associated with diseases affecting cognitive capacities. *DYRK1A*, which lies in the Down syndrome critical region, is thought to underlie some of the cognitive impairment associated with having three copies of chromosome 21 (64). Mutations in *NRG3* have been associated with schizophrenia, a condition that has been suggested to affect human-specific cognitive traits (65, 66). Mutations in *CADPS2* have been implicated in autism (67), as have mutations in *AUTS2* (68). Autism is a developmental disorder of brain function in which social interactions, communication, activity, and interest patterns are affected, as well as cognitive aspects crucial for human sociality and culture

(69). It may thus be that multiple genes involved in cognitive development were positively selected during the early history of modern humans.

One gene of interest may be *RUNX2* (*CBFA1*). It is the only gene in the genome known to cause cleidocranial dysplasia, which is characterized by delayed closure of cranial sutures, hypoplastic or aplastic clavicles, a bell-shaped rib cage, and dental abnormalities (70). Some of these features affect morphological traits for which modern humans differ from Neandertals as well as other earlier hominins. For example, the cranial malformations seen in cleidocranial dysplasia include frontal bossing, i.e., a protruding frontal bone. A more prominent frontal bone is a feature that differs between modern humans and Neandertals as well as other archaic hominins. The clavicle, which is affected in cleidocranial dysplasia, differs in morphology between modern humans and Neandertals (71) and is associated with a different architecture of the shoulder joint. Finally, a bell-shaped rib cage is typical of Neandertals and other archaic hominins. A reasonable hypothesis is thus that an evolutionary change in *RUNX2* was of importance in the origin of modern humans and that this change affected aspects of the morphology of the upper body and cranium.

Population divergence of Neandertals and modern humans. A long-standing question is when the ancestral populations of Neandertals and modern humans diverged. Population divergence, defined as the time point when two populations last exchanged genes, is more recent than the DNA sequence divergence because the latter is the sum of the time to population divergence plus the average time to the common ancestors of DNA sequences within the ancestral population. The divergence time of two populations can be

Table 3. Top 20 candidate selective sweep regions.

Region (hg18)	S	Width (cM)	Gene(s)
chr2:43265008-43601389	-6.04	0.5726	<i>ZFP36L2;THADA</i>
chr11:95533088-95867597	-4.78	0.5538	<i>JRKL;CCDC82;MAML2</i>
chr10:62343313-62655667	-6.1	0.5167	<i>RHOBTB1</i>
chr21:37580123-37789088	-4.5	0.4977	<i>DYRK1A</i>
chr10:83336607-83714543	-6.13	0.4654	<i>NRG3</i>
chr14:100248177-100417724	-4.84	0.4533	<i>MIR337;MIR665;DLK1;RTL1;MIR431;MIR493;MEG3;MIR770</i>
chr3:157244328-157597592	-6	0.425	<i>KCNAB1</i>
chr11:30601000-30992792	-5.29	0.3951	
chr2:176635412-176978762	-5.86	0.3481	<i>HOXD11;HOXD8;EVX2;MTX2;HOXD1;HOXD10;HOXD13;HOXD4;HOXD12;HOXD9;MIR10B;HOXD3</i>
chr11:71572763-71914957	-5.28	0.3402	<i>CLPB;FOLR1;PHOX2A;FOLR2;INPPL1</i>
chr7:41537742-41838097	-6.62	0.3129	<i>INHBA</i>
chr10:60015775-60262822	-4.66	0.3129	<i>BICC1</i>
chr6:45440283-45705503	-4.74	0.3112	<i>RUNX2;SUPT3H</i>
chr1:149553200-149878507	-5.69	0.3047	<i>SELENBP1;POGZ;MIR554;RFX5;SNX27;CGN;TUFT1;PI4KB;PSMB4</i>
chr7:121763417-122282663	-6.35	0.2855	<i>RNF148;RNF133;CADPS2</i>
chr7:93597127-93823574	-5.49	0.2769	
chr16:62369107-62675247	-5.18	0.2728	
chr14:48931401-49095338	-4.53	0.2582	
chr6:90762790-90903925	-4.43	0.2502	<i>BACH2</i>
chr10:9650088-9786954	-4.56	0.2475	

inferred from the frequency with which derived alleles of SNPs discovered in one population are seen in the other population. The reason for this is that the older the population divergence, the more likely it is that derived alleles discovered in one population are due to novel mutations in that population. We compared transversion SNPs identified in a Yoruba individual (33) to other humans and used the chimpanzee and orangutan genomes to identify the ancestral alleles. We found that the proportion of derived alleles is 30.6% in the Yoruba, 29.8% in the Han Chinese, 29.7% in the French, 29.3% in the Papuan, 26.3% in the San, and 18.0% in Neandertals. We used four models of Yoruba demographic history to translate derived allele fractions to population divergence (SOM Text 14). All provided similar estimates. Assuming that human-chimpanzee average DNA sequence divergence was 5.6 to 8.3 million years ago, this suggests that Neandertals and present-day human populations separated between 270,000 and 440,000 years ago (SOM Text 14), a date that is compatible with some interpretations of the paleontological and archaeological record (2, 72).

Neandertals are closer to non-Africans than to Africans. To test whether Neandertals are more closely related to some present-day humans than to others, we identified SNPs by comparing one randomly chosen sequence from each of two present-day humans and asking if the Neandertals match the alleles of the two individuals equally often. If gene flow between Neandertals and modern humans ceased before differentiation between present-day human populations began, this is expected to be the case no matter which present-day humans are compared. The prediction of this null hypothesis of no gene flow holds regardless of population expansions, bottlenecks, or substructure that might have occurred in modern human history (SOM Text 15). The reason for this is that when single chromosomes are analyzed in the two present-day populations, differences in demographic histories in the two populations will not affect the results even if they may profoundly influence allele frequencies. Under the alternative model of later gene flow between Neandertals and modern humans, we expect Neandertals to match alleles in individuals from some parts of the world more often than the others.

We restricted this analysis to biallelic SNPs where two present-day humans carry different alleles and where the Neandertals carried the derived allele, i.e., not matching chimpanzee. We measured the difference in the percent matching by a statistic $D(H_1, H_2, \text{Neandertal}, \text{chimpanzee})$ (SOM Text 15) that does not differ significantly from zero when the derived alleles in the Neandertal match alleles in the two humans equally often. If D is positive, Neandertal alleles match alleles in the second human (H_2) more often, while if D is negative, Neandertal alleles match alleles in the first human (H_1) more often. We performed this test using eight present-day humans: two European Americans (CEU), two East Asians

(ASN), and four West Africans (YRI), for whom sequences have been generated with Sanger technology, with reads of ~750 bp that we mapped along with the Neandertal reads to the chimpanzee genome. We find that the Neandertals are equally close to Europeans and East Asians: $D(\text{ASN}, \text{CEU}, \text{Neandertal}, \text{chimpanzee}) = -0.53 \pm 0.46\%$ (<1.2 SD from 0% or $P = 0.25$). However, the Neandertals are significantly closer to non-Africans than to Africans: $D(\text{YRI}, \text{CEU}, \text{Neandertal}, \text{chimpanzee}) = 4.57 \pm 0.39\%$ and $D(\text{YRI}, \text{ASN}, \text{Neandertal}, \text{chimpanzee}) = 4.81 \pm 0.39\%$ (both >11 SD from 0% or $P < 10^{-12}$) (table S51).

The greater genetic proximity of Neandertals to Europeans and Asians than to Africans is seen no matter how we subdivide the data: (i) by individual pairs of humans (Table 4), (ii) by chromosome, (iii) by substitutions that are transitions or transversions, (iv) by hypermutable CpG versus all other sites, (v) by Neandertal sequences shorter or longer than 50 bp, and (vi) by 454 or Illumina data. It is also seen when we restrict the analysis to A/T and C/G substitutions, showing that our observations are unlikely to be due to biased allele calling or biased gene conversion (SOM Text 15).

A potential artifact that might explain these observations is contamination of the Neandertal sequences with non-African DNA. However, the magnitude of contamination necessary to explain the CEU-YRI and ASN-YRI comparisons are both over 10% and thus inconsistent with our estimates of contamination in the Neandertal data, which are all below 1% (Table 1). In addition to the low estimates of contamination, there are two reasons that contamination cannot explain our results. First, when we analyze the three Neandertal bones Vi33.16, Vi33.25, and Vi33.26 separately, we obtain consistent values of the D statistics, which is unlikely to arise under the hypothesis of contamination because each specimen was individually handled and was thus unlikely to have been affected by the same degree of contamination (SOM Text 15). Second, if European contamination explains the skews, the ratio $D(H_1, H_2, \text{Neandertal}, \text{chimpanzee})/D(H_1, H_2, \text{European}, \text{chimpanzee})$ should provide a direct estimate of the contamination proportion α , because the ratio measures how close the Neandertal data are to what would be expected from entirely European contamination. However, when we estimate α for all three population pairs, we obtain statistically inconsistent results: $\alpha = 13.9 \pm 1.1\%$ for $H_1-H_2 = \text{CEU-YRI}$, $\alpha = 18.9 \pm 1.9\%$ for ASN-YRI , and $\alpha = -3.9 \pm 5.1\%$ for CEU-ASN . This indicates that the skews cannot be explained by a unifying hypothesis of European contamination.

To analyze the relationship of the Neandertals to a more diverse set of modern humans, we repeated the analysis above using the genome sequences of the French, Han, Papuan, Yoruba, and San individuals that we generated (SOM Text 9). Strikingly, no comparison within Eurasia (Papuan-French-Han) or within Africa (Yoruba-San) shows significant skews in D ($|Z| < 2$ SD).

However, all comparisons of non-Africans and Africans show that the Neandertal is closer to the non-African (D from 3.8% to 5.3%, $|Z| > 7.0$ SD) (Table 4). Thus, analyses of present-day humans consistently show that Neandertals share significantly more derived alleles with non-Africans than with Africans, whereas they share equal amounts of derived alleles when compared either to individuals within Eurasia or to individuals within Africa.

Direction of gene flow. A parsimonious explanation for these observations is that Neandertals exchanged genes with the ancestors of non-Africans. To determine the direction of gene flow consistent with the data, we took advantage of the fact that non-Africans are more distantly related to San than to Yoruba (73–75) (Table 4). This is reflected in the fact that $D(P, \text{San}, Q, \text{chimpanzee})$ is 1.47 to 1.68 times greater than $D(P, \text{Yoruba}, Q, \text{chimpanzee})$, where P and Q are non-Africans (SOM Text 15). Under the hypothesis of modern human to Neandertal gene flow, $D(P, \text{San}, \text{Neandertal}, \text{chimpanzee})$ should be greater than $D(P, \text{Yoruba}, \text{Neandertal}, \text{chimpanzee})$ by the same amount, because the deviation of the D statistics is due to Neandertals inheriting a proportion of ancestry from a non-African-like population Q. Empirically, however, the ratio is significantly smaller (1.00 to 1.03, $P < 0.0002$) (SOM Text 15). Thus, all or almost all of the gene flow detected was from Neandertals into modern humans.

Segments of Neandertal ancestry in non-African genomes. If Neandertal-to-modern human gene flow occurred, we predict that we should find DNA segments with an unusually low divergence to Neandertal in present-day humans. Furthermore, we expect that such segments will tend to have an unusually high divergence to other present-day humans because they come from Neandertals. In the absence of gene flow, segments with low divergence to Neandertals are expected to arise due to other effects, for example, a low mutation rate in a genomic segment since the split from the chimpanzee lineage. However, this will cause present-day humans to tend to have low divergence from each other in such segments, i.e., the opposite effect from gene flow. The qualitative distinction between these predictions allows us to detect a signal of gene flow. To search for segments with relatively few differences between Neandertals and present-day humans, we used haploid human DNA sequences, because in a diploid individual, both alleles would have to be derived from Neandertals to produce a strong signal. To obtain haploid human sequences, we took advantage of the fact that the human genome reference sequence is composed of a tiling path of bacterial artificial chromosomes (BACs), which each represent single human haplotypes over scales of 50 to 150 kb, and we focused on BACs from RPC111, the individual that contributed about two-thirds of the reference sequence and that has been previously shown to be of about 50% European and 50% African ancestry (SOM Text 16) (76). We then estimated the Neandertal to present-

day human divergence and found that in the extreme tail of low-divergence BACs there was a greater proportion of European segments than African segments, consistent with the notion that some genomic segments (SOM Text 16) were exchanged between Neandertals and non-Africans.

To determine whether these segments are unusual in their divergence to other present-day humans, we examined the divergence of each segment to the genome of Craig Venter (77). We find that present-day African segments with the lowest divergence to Neandertals have a divergence to Venter that is 35% of the genome-wide average and that their divergence to Venter in-

creases monotonically with divergence to Neandertals, as would be expected if these segments were similar in Neandertals and present-day humans due to, for example, a low mutation rate in these segments (Fig. 5A). In contrast, the European segments with the lowest divergence to Neandertals have a divergence to Venter that is 140% of the genome-wide average, which drops precipitously with increasing divergence to humans before rising again (Fig. 5A). This nonmonotonic behavior is significant at $P < 10^{-9}$ and is unexpected in the absence of gene flow from Neandertals into the ancestors of non-Africans. The reason for this is that other causes for a low di-

vergence to Neandertals, such as low mutation rates, contamination by modern non-African DNA, or gene flow into Neandertals, would produce monotonic behaviors. Among the segments with low divergence to Neandertals and high divergence to Venter, 94% of segments are of European ancestry (Fig. 5B), suggesting that segments of likely Neandertal ancestry in present-day humans can be identified with relatively high confidence.

Non-Africans haplotypes match Neandertals unexpectedly often. An alternative approach to detect gene flow from Neandertals into modern humans is to focus on patterns of variation in present-day humans—blinded to information from

Table 4. Neandertals are more closely related to present-day non-Africans than to Africans. For each pair of modern humans H_1 and H_2 that we examined, we reported D (H_1 , H_2 , Neandertal, Chimpanzee): the difference in the percentage matching of Neandertal to two humans at sites where Neandertal does not match chimpanzee, with ± 1 standard

error. Values that deviate significantly from 0% after correcting for 38 hypotheses tested are highlighted in bold ($|Z| > 2.8$ SD). Neandertal is skewed toward matching non-Africans more than Africans for all pairwise comparisons. Comparisons within Africans or within non-Africans are all consistent with 0%.

Population comparison	H_1	H_2	% Neandertal matching to H_2 – % Neandertal matching to H_1 (± 1 standard error)
<i>ABI3730 sequencing (~750 bp reads) used to discover H_1-H_2 differences</i>			
African to African	NA18517 (Yoruba)	NA18507 (Yoruba)	-0.1 \pm 0.6
	NA18517 (Yoruba)	NA19240 (Yoruba)	1.5 \pm 0.7
	NA18517 (Yoruba)	NA19129 (Yoruba)	-0.1 \pm 0.7
	NA18507 (Yoruba)	NA19240 (Yoruba)	-0.5 \pm 0.6
	NA18507 (Yoruba)	NA19129 (Yoruba)	0.0 \pm 0.5
	NA19240 (Yoruba)	NA19129 (Yoruba)	-0.6 \pm 0.7
African to Non-African	NA18517 (Yoruba)	NA12878 (European)	4.1 \pm 0.8
	NA18517 (Yoruba)	NA12156 (European)	5.1 \pm 0.7
	NA18517 (Yoruba)	NA18956 (Japanese)	2.9 \pm 0.8
	NA18517 (Yoruba)	NA18555 (Chinese)	3.9 \pm 0.7
	NA18507 (Yoruba)	NA12878 (European)	4.2 \pm 0.6
	NA18507 (Yoruba)	NA12156 (European)	5.5 \pm 0.6
	NA18507 (Yoruba)	NA18956 (Japanese)	5.0 \pm 0.7
	NA18507 (Yoruba)	NA18555 (Chinese)	5.8 \pm 0.6
	NA19240 (Yoruba)	NA12878 (European)	3.5 \pm 0.7
	NA19240 (Yoruba)	NA12156 (European)	3.1 \pm 0.7
	NA19240 (Yoruba)	NA18956 (Japanese)	2.7 \pm 0.7
	NA19240 (Yoruba)	NA18555 (Chinese)	5.4 \pm 0.9
	NA19129 (Yoruba)	NA12878 (European)	3.9 \pm 0.7
	NA19129 (Yoruba)	NA12156 (European)	4.9 \pm 0.7
	NA19129 (Yoruba)	NA18956 (Japanese)	5.1 \pm 0.8
Non-African to Non-African	NA19129 (Yoruba)	NA18555 (Chinese)	4.7 \pm 0.8
	NA12878 (European)	NA12156 (European)	-0.5 \pm 0.8
	NA12878 (European)	NA18956 (Japanese)	0.4 \pm 0.8
	NA12878 (European)	NA18555 (Chinese)	0.3 \pm 0.8
	NA12156 (European)	NA18956 (Japanese)	-0.3 \pm 0.8
	NA12156 (European)	NA18555 (Chinese)	1.3 \pm 0.7
<i>Illumina GAI sequencing (~76 bp reads) used to discover H_1-H_2 differences</i>			
African - African	HGDP01029 (San)	HGDP01029 (Yoruba)	-0.1 \pm 0.4
African to Non-African	HGDP01029 (San)	HGDP00521 (French)	4.2 \pm 0.4
	HGDP01029 (San)	HGDP00542 (Papuan)	3.9 \pm 0.5
	HGDP01029 (San)	HGDP00778 (Han)	5.0 \pm 0.5
	HGDP01029 (Yoruba)	HGDP00521 (French)	4.5 \pm 0.4
	HGDP01029 (Yoruba)	HGDP00542 (Papuan)	4.4 \pm 0.6
	HGDP01029 (Yoruba)	HGDP00778 (Han)	5.3 \pm 0.5
Non-African to Non-African	HGDP00521 (French)	HGDP00542 (Papuan)	0.1 \pm 0.5
	HGDP00521 (French)	HGDP00778 (Han)	1.0 \pm 0.6
	HGDP00542 (Papuan)	HGDP00778 (Han)	0.7 \pm 0.6

the Neandertal genome—in order to identify regions that are the strongest candidates for being derived from Neandertals. If these candidate regions match the Neandertals at a higher rate than is expected by chance, this provides additional evidence for gene flow from Neandertals into modern humans.

We thus identified regions in which there is considerably more diversity outside Africa than

inside Africa, as might be expected in regions that have experienced gene flow from Neandertals to non-Africans. We used 1,263,750 Perlegen Class A SNPs, identified in individuals of diverse ancestry (78), and found 13 candidate regions of Neandertal ancestry (SOM Text 17). A prediction of Neandertal-to-modern human gene flow is that DNA sequences that entered the human gene pool from Neandertals will tend to match Neandertal

more often than their frequency in the present-day human population. To test this prediction, we identified 166 “tag SNPs” that separate 12 of the haplotype clades in non-Africans (OOA) from the cosmopolitan haplotype clades shared between Africans and non-Africans (COS) and for which we had data from the Neandertals. Overall, the Neandertals match the deep clade unique to non-Africans at 133 of the 166 tag SNPs, and 10 of the

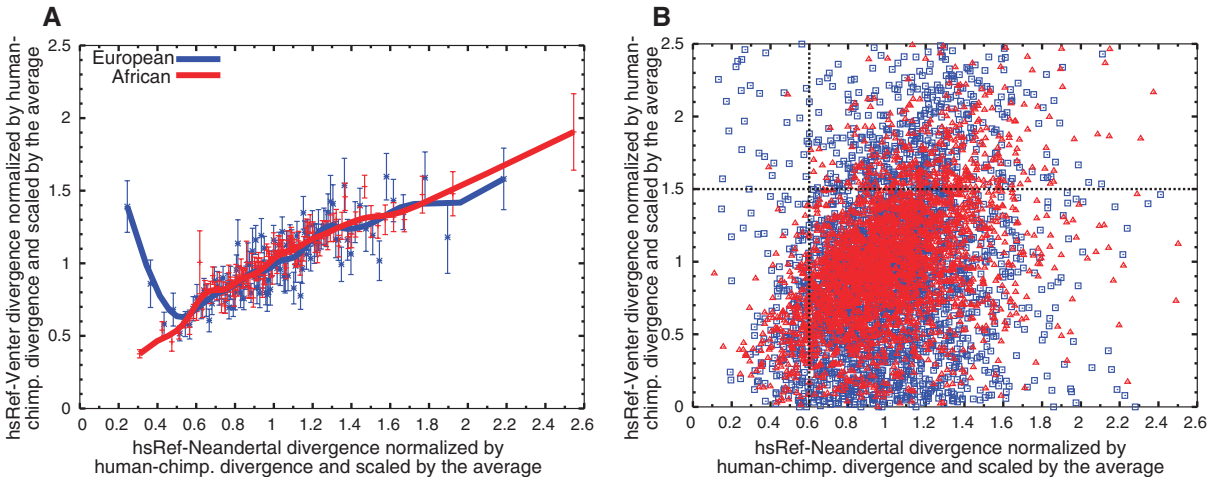


Fig. 5. Segments of Neandertal ancestry in the human reference genome. We examined 2825 segments in the human reference genome that are of African ancestry and 2797 that are of European ancestry. **(A)** European segments, with few differences from the Neandertals, tend to have many differences from other present-day humans, whereas African segments do

not, as expected if the former are derived from Neandertals. **(B)** Scatter plot of the segments in **(A)** with respect to their divergence to the Neandertals and to Venter. In the top left quadrant, 94% of segments are of European ancestry, suggesting that many of them are due to gene flow from Neandertals.

Table 5. Non-African haplotypes match Neandertal at an unexpected rate. We identified 13 candidate gene flow regions by using 48 CEU+ASN to represent the OOA population, and 23 African Americans to represent the AFR population. We identified tag SNPs for each region that separate an out-of-Africa specific clade (OOA) from a cosmopolitan clade (COS) and then assessed the rate at

which Neandertal matches each of these clades by further subdividing tag SNPs based on their ancestral and derived status in Neandertal and whether they match the OOA-specific clade or not. Thus, the categories are AM (Ancestral Nonmatch), DN (Derived Nonmatch), DM (Derived Match), and AN (Ancestral Match). We do not list the sites where matching is ambiguous.

Chromosome	Start of candidate region in Build 36	End of candidate region in Build 36	Span (bp)	S_T (estimated ratio of OOA/AFR gene tree depth)	Average frequency of tag in OOA clade	Neandertal (M)atches OOA-specific clade AM DM		Neandertal does (N)ot match OOA-specific clade AN DN		Qualitative assessment*
1	168,110,000	168,220,000	110,000	2.9	6.3%	5	10	1	0	OOA
1	223,760,000	223,910,000	150,000	2.8	6.3%	1	4	0	0	OOA
4	171,180,000	171,280,000	100,000	1.9	5.2%	1	2	0	0	OOA
5	28,950,000	29,070,000	120,000	3.8	3.1%	16	16	6	0	OOA
6	66,160,000	66,260,000	100,000	5.7	28.1%	6	6	0	0	OOA
9	32,940,000	33,040,000	100,000	2.8	4.2%	7	14	0	0	OOA
10	4,820,000	4,920,000	100,000	2.6	9.4%	9	5	0	0	OOA
10	38,000,000	38,160,000	160,000	3.5	8.3%	5	9	2	0	OOA
10	69,630,000	69,740,000	110,000	4.2	19.8%	2	2	0	1	OOA
15	45,250,000	45,350,000	100,000	2.5	1.1%	5	6	1	0	OOA
17	35,500,000	35,600,000	100,000	2.9	(no tags)	—	—	—	—	—
20	20,030,000	20,140,000	110,000	5.1	64.6%	0	0	10	5	COS
22	30,690,000	30,820,000	130,000	3.5	4.2%	0	2	5	2	COS
Relative tag SNP frequencies in actual data						34%	46%	15%	5%	
Relative tag SNP simulated under a demographic model without introgression						34%	5%	33%	27%	
Relative tag SNP simulated under a demographic model with introgression						23%	31%	37%	9%	

*To qualitatively assess the regions in terms of which clade the Neandertal matches, we asked whether the proportion matching the OOA-specific clade (AM and DM) is much more than 50%. If so, we classify it as an OOA region, and otherwise a COS region. One region is unclassified because no tag SNPs were found. We also compared to simulations with and without gene flow (SOM Text 17), which show that the rate of DM and DN tag SNPs where Neandertal is derived are most informative for distinguishing gene flow from no gene flow.

12 regions where tag SNPs occur show an excess of OOA over COS sites. Given that the OOA alleles occur at a frequency of much less than 50% in non-Africans (average of 13%, and all less than 30%) (Table 5), the fact that the candidate regions match the Neandertals in 10 of 12 cases ($P = 0.019$) suggests that they largely derive from Neandertals. The proportion of matches is also larger than can be explained by contamination, even if all Neandertal data were composed of present-day non-African DNA ($P = 0.0025$) (SOM Text 17).

This analysis shows that some old haplotypes most likely owe their presence in present-day non-Africans to gene flow from Neandertals. However, not all old haplotypes in non-Africans may have such an origin. For example, it has been suggested that the H2 haplotype on chromosome 17 and the D haplotype of the microcephalin gene were contributed by Neandertals to present-day non-Africans (12, 79, 80). This is not supported by the current data because the Neandertals analyzed do not carry these haplotypes.

The extent of Neandertal ancestry. To estimate the proportion of Neandertal ancestry, we

compare the similarity of non-Africans to Neandertals with the similarity of two Neandertals, N1 and N2, to each other. Under the assumption that there was no gene flow from Neandertals to the ancestors of modern Africans, the proportion of Neandertal ancestry of non-Africans, f , can be estimated by the ratio $S(OOA, AFR, N1, Chimpanzee) / S(N2, AFR, N1, Chimpanzee)$, where the S statistic is an unnormalized version of the D statistic (SOM Text 18, Eq. S18.4). Using Neandertals from Vindija, as well as Mezmaiskaya, we estimate f to be between 1.3% and 2.7% (SOM Text 18). To obtain an independent estimate of f , we fit a population genetic model to the D statistics in Table 4 and SOM Text 15 as well as to other summary statistics of the data. Assuming that gene flow from Neandertals occurred between 50,000 and 80,000 years ago, this method estimates f to be between 1 and 4%, consistent with the above estimate (SOM Text 19). We note that a previous study found a pattern of genetic variation in present-day humans that was hypothesized to be due to gene flow from Neandertals or other archaic hominins into modern humans (81). The authors of this study estimated the fraction of non-African genomes affected by “archaic” gene flow to be 14%, almost an order of magnitude greater than our estimates, suggesting that their observations may not be entirely explained by gene flow from Neandertals.

Implications for modern human origins.

One model for modern human origins suggests that all present-day humans trace all their ancestry back to a small African population that expanded and replaced archaic forms of humans without admixture. Our analysis of the Neandertal genome may not be compatible with this view because Neandertals are on average closer to individuals in Eurasia than to individuals in Africa. Furthermore, individuals in Eurasia today carry regions in their genome that are closely related to those in Neandertals and distant from other present-day humans. The data suggest that between 1 and 4% of the genomes of people in Eurasia are derived from Neandertals. Thus, while the Neandertal genome presents a challenge to the simplest version of an “out-of-Africa” model for modern human origins, it continues to support the view that the vast majority of genetic variants that exist at appreciable frequencies outside Africa came from Africa with the spread of anatomically modern humans.

A striking observation is that Neandertals are as closely related to a Chinese and Papuan individual as to a French individual, even though morphologically recognizable Neandertals exist only in the fossil record of Europe and western Asia. Thus, the gene flow between Neandertals and modern humans that we detect most likely occurred before the divergence of Europeans, East Asians, and Papuans. This may be explained by mixing of early modern humans ancestral to present-day non-Africans with Neandertals in the Middle East before their expansion into Eurasia. Such a scenario is compatible with the archaeo-

logical record, which shows that modern humans appeared in the Middle East before 100,000 years ago whereas the Neandertals existed in the same region after this time, probably until 50,000 years ago (82).

It is important to note that although we detect a signal compatible with gene flow from Neandertals into ancestors of present-day humans outside Africa, this does not show that other forms of gene flow did not occur (Fig. 6). For example, we detect gene flow from Neandertals into modern humans but no reciprocal gene flow from modern humans into Neandertals. Although gene flow between different populations need not be bidirectional, it has been shown that when a colonizing population (such as anatomically modern humans) encounters a resident population (such as Neandertals), even a small number of breeding events along the wave front of expansion into new territory can result in substantial introduction of genes into the colonizing population as introduced alleles can “surf” to high frequency as the population expands. As a consequence, detectable gene flow is predicted to almost always be from the resident population into the colonizing population, even if gene flow also occurred in the other direction (83). Another prediction of such a surfing model is that even a very small number of events of interbreeding can result in appreciable allele frequencies of Neandertal alleles in the present-day populations. Thus, the actual amount of interbreeding between Neandertals and modern humans may have been very limited, given that it contributed only 1 to 4% of the genome of present-day non-Africans.

It may seem surprising that we see no evidence for greater gene flow from Neandertals to present-day Europeans than to present-day people in eastern Asia given that the morphology of some hominin fossils in Europe has been interpreted as evidence for gene flow from Neandertals into early modern humans late in Neandertal history [e.g., (84)] (Fig. 6). It is possible that later migrations into Europe, for example in connection with the spread of agriculture, have obscured the traces of such gene flow. This possibility can be addressed by the determination of genome sequences from preagricultural early modern humans in Europe (85). It is also possible that if the expansion of modern humans occurred differently in Europe than in the Middle East, for example by already large populations interacting with Neandertals, then there may be little or no trace of any gene flow in present-day Europeans even if interbreeding occurred. Thus, the contingencies of demographic history may cause some events of past interbreeding to leave traces in present-day populations, whereas other events will leave little or no traces. Obviously, gene flow that left little or no traces in the present-day gene pool is of little or no consequence from a genetic perspective, although it may be of interest from a historical perspective.

Although gene flow from Neandertals into modern humans when they first left sub-Saharan Africa seems to be the most parsimonious model

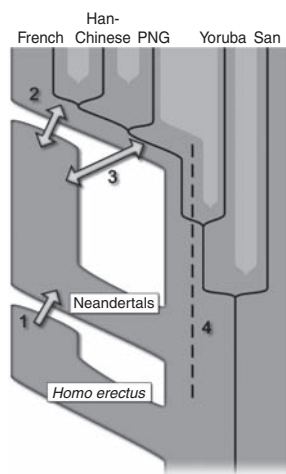


Fig. 6. Four possible scenarios of genetic mixture involving Neandertals. Scenario 1 represents gene flow into Neandertal from other archaic hominins, here collectively referred to as *Homo erectus*. This would manifest itself as segments of the Neandertal genome with unexpectedly high divergence from present-day humans. Scenario 2 represents gene flow between late Neandertals and early modern humans in Europe and/or western Asia. We see no evidence of this because Neandertals are equally distantly related to all non-Africans. However, such gene flow may have taken place without leaving traces in the present-day gene pool. Scenario 3 represents gene flow between Neandertals and the ancestors of all non-Africans. This is the most parsimonious explanation of our observation. Although we detect gene flow only from Neandertals into modern humans, gene flow in the reverse direction may also have occurred. Scenario 4 represents old substructure in Africa that persisted from the origin of Neandertals until the ancestors of non-Africans left Africa. This scenario is also compatible with the current data.

compatible with the current data, other scenarios are also possible. For example, we cannot currently rule out a scenario in which the ancestral population of present-day non-Africans was more closely related to Neandertals than the ancestral population of present-day Africans due to ancient substructure within Africa (Fig. 6). If after the divergence of Neandertals there was incomplete genetic homogenization between what were to become the ancestors of non-Africans and Africans, present-day non-Africans would be more closely related to Neandertals than are Africans. In fact, old population substructure in Africa has been suggested based on genetic (81) as well as paleontological data (86).

In conclusion, we show that genome sequences from an extinct late Pleistocene hominin can be reliably recovered. The analysis of the Neandertal genome shows that they are likely to have had a role in the genetic ancestry of present-day humans outside of Africa, although this role was relatively minor given that only a few percent of the genomes of present-day people outside Africa are derived from Neandertals. Our results also point to a number of genomic regions and genes as candidates for positive selection early in modern human history, for example, those involved in cognitive abilities and cranial morphology. We expect that further analyses of the Neandertal genome as well as the genomes of other archaic hominins will generate additional hypotheses and provide further insights into the origins and early history of present-day humans.

References and Notes

1. J. L. Bischoff *et al.*, *High-Resolution U-Series Dates from the Sima de los Huesos Hominids Yields 600+/-66 kys: Implications for the Evolution of the Early Neanderthal Lineage* (Elsevier, Amsterdam, PAYS-BAS, 2007), vol. 34.
2. J. J. Hublin, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 16022 (2009).
3. C. B. Stringer, J. Hublin, *J. Hum. Evol.* **37**, 873 (1999).
4. C. Finlayson *et al.*, *Nature* **443**, 850 (2006).
5. J. Krause *et al.*, *Nature* **449**, 902 (2007).
6. R. Grün *et al.*, *J. Hum. Evol.* **49**, 316 (2005).
7. N. Mercier, H. Valladas, in *Late Quaternary Chronology and Palaeoclimate of the Eastern Mediterranean, Radiocarbon*, O. Bar-Yosef, R. Kra, Eds. (1994), pp. 13–20.
8. E. Trinkaus *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 11231 (2003).
9. J. Zilhão, E. Trinkaus, in *Trabalhos de Arqueologia* (Instituto Português de Arqueologia, Lisbon, 2002), vol. 22.
10. S. E. Bailey, T. D. Weaver, J. J. Hublin, *J. Hum. Evol.* **57**, 11 (2009).
11. G. Bräuer, H. Broeg, C. Stringer, in *Neanderthals Revisited: New Approaches and Perspectives*. (2006), pp. 269–279.
12. P. D. Evans, N. Mekel-Bobrov, E. J. Vallender, R. R. Hudson, B. T. Lahn, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18178 (2006).
13. J. D. Wall, M. F. Hammer, *Curr. Opin. Genet. Dev.* **16**, 606 (2006).
14. M. Currat, L. Excoffier, *PLoS Biol.* **2**, e421 (2004).
15. A. W. Briggs *et al.*, *Science* **325**, 318 (2009).
16. M. Krings *et al.*, *Cell* **90**, 19 (1997).
17. L. Orlando *et al.*, *Curr. Biol.* **16**, R400 (2006).
18. I. V. Ovchinnikov *et al.*, *Nature* **404**, 490 (2000).
19. D. Serre *et al.*, *PLoS Biol.* **2**, E57 (2004).
20. S. Pääbo, *Trends Cell Biol.* **9**, M13 (1999).
21. S. Pääbo, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1939 (1989).
22. S. Pääbo *et al.*, *Annu. Rev. Genet.* **38**, 645 (2004).
23. A. W. Briggs *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14616 (2007).
24. P. Brotherton *et al.*, *Nucleic Acids Res.* **35**, 5717 (2007).
25. M. Hofreiter, V. Jaenicke, D. Serre, A. von Haeseler, S. Pääbo, *Nucleic Acids Res.* **29**, 4793 (2001).
26. M. Höss, P. Jaruga, T. H. Zastawny, M. Dizdargolu, S. Pääbo, *Nucleic Acids Res.* **24**, 1304 (1996).
27. R. K. Saiki *et al.*, *Science* **230**, 1350 (1985).
28. C. Lalueza-Fox *et al.*, *Science* **318**, 1453 (2007).
29. J. Krause *et al.*, *Curr. Biol.* **17**, 1908 (2007).
30. C. Lalueza-Fox *et al.*, *BMC Evol. Biol.* **8**, 342 (2008).
31. C. Lalueza-Fox, E. Gigli, M. de la Rasilla, J. Fortea, A. Rosas, *Biol. Lett.* **5**, 809 (2009).
32. J. Krause *et al.*, *Nature* **439**, 724 (2006).
33. D. R. Bentley *et al.*, *Nature* **456**, 53 (2008).
34. M. Margulies *et al.*, *Nature* **437**, 376 (2005).
35. H. N. Poinar *et al.*, *Science* **311**, 392 (2006).
36. M. Rasmussen *et al.*, *Nature* **463**, 757 (2010).
37. M. Stiller *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13578 (2006).
38. W. Miller *et al.*, *Nature* **456**, 387 (2008).
39. K. Prüfer *et al.*, *Genome Biol.* **11**, R47 (2010).
40. R. E. Green *et al.*, *Nature* **444**, 330 (2006).
41. R. E. Green *et al.*, *EMBO J.* **28**, 2494 (2009).
42. J. P. Noonan *et al.*, *Science* **314**, 1113 (2006).
43. A. D. Greenwood, C. Capelli, G. Possnert, S. Pääbo, *Mol. Biol. Evol.* **16**, 1466 (1999).
44. J. D. Wall, S. K. Kim, *PLoS Genet.* **3**, e175 (2007).
45. R. E. Green *et al.*, *Cell* **134**, 416 (2008).
46. A. W. Briggs *et al.*, *J. Vis. Exp.* **2009**, 1573 (2009).
47. T. Maricic, S. Pääbo, *Biotechniques* **46**, 51, 54 (2009).
48. M. Kircher, U. Stenzel, J. Kelso, *Genome Biol.* **10**, R83 (2009).
49. A. W. Briggs *et al.*, *Nucleic Acids Res.* **38**, e87 (2010).
50. J. C. Dohm, C. Lottaz, T. Borodina, H. Himmelbauer, *Nucleic Acids Res.* **36**, e105 (2008).
51. B. Paten *et al.*, *Genome Res.* **18**, 1829 (2008).
52. M. Goodman, *Am. J. Hum. Genet.* **64**, 31 (1999).
53. T. de Torres *et al.*, *Archaeometry* published online 29 October 2009; 10.1111/j.1475-4754.2009.00491.x.
54. R. W. Schmitz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13342 (2002).
55. A. R. Skinner *et al.*, *Appl. Radiat. Isot.* **62**, 219 (2005).
56. N. Patterson, D. J. Richter, S. Gnerre, E. S. Lander, D. Reich, *Nature* **441**, 1103 (2006).
57. H. A. Burbano *et al.*, *Science* **328**, 723 (2010).
58. Z. Zhang *et al.*, *Mol. Cell. Proteomics* **4**, 914 (2005).
59. N. Matsuyoshi, S. Imamura, *Biochem. Biophys. Res. Commun.* **235**, 355 (1997).
60. P. Richard, J. L. Manley, *Genes Dev.* **23**, 1247 (2009).
61. M. Huber *et al.*, *J. Invest. Dermatol.* **124**, 998 (2005).
62. C. Alkan *et al.*, *Nat. Genet.* **41**, 1061 (2009).
63. H. Parikh, V. Lyssenko, L. C. Groop, *BMC Med. Genomics* **2**, 72 (2009).
64. B. Hämmerle, C. Elizalde, J. Galceran, W. Becker, F. J. Tejedor, *J. Neural Transm. Suppl.* **2003**, 129 (2003).
65. T. J. Crow, *Eur. Neuropsychopharmacol.* **5** (suppl), 59 (1995).
66. P. Khaitovich *et al.*, *Genome Biol.* **9**, R124 (2008).
67. T. Sadakata *et al.*, *J. Clin. Invest.* **117**, 931 (2007).
68. R. Sultana *et al.*, *Genomics* **80**, 129 (2002).
69. M. Tomasello, M. Carpenter, J. Call, T. Behne, H. Moll, *Behav. Brain Sci.* **28**, 675, discussion 691 (2005).
70. S. Mundlos *et al.*, *Cell* **89**, 773 (1997).
71. J. L. Voisin, *J. Hum. Evol.* **55**, 438 (2008).
72. T. D. Weaver, C. C. Roseman, C. B. Stringer, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4645 (2008).
73. D. M. Behar *et al.*, Genographic Consortium, *Am. J. Hum. Genet.* **82**, 1130 (2008).
74. J. X. Sun, J. C. Mullikin, N. Patterson, D. E. Reich, *Mol. Biol. Evol.* **26**, 1017 (2009).
75. E. T. Wood *et al.*, *Eur. J. Hum. Genet.* **13**, 867 (2005).
76. D. Reich *et al.*, *PLoS Genet.* **5**, e1000360 (2009).
77. S. Levy *et al.*, *PLoS Biol.* **5**, e254 (2007).
78. D. A. Hinds *et al.*, *Science* **307**, 1072 (2005).
79. J. Hardy *et al.*, *Biochem. Soc. Trans.* **33**, 582 (2005).
80. H. Stefansson *et al.*, *Nat. Genet.* **37**, 129 (2005).
81. J. D. Wall, K. E. Lohmueller, V. Plagnol, *Mol. Biol. Evol.* **26**, 1823 (2009).
82. O. Bar-Yosef, in *Neandertals and Modern Humans in Western Asia*, T. Akazawa, K. Aoki, O. Bar-Yosef, Eds. (Plenum, New York, 1999), pp. 39–56.
83. M. Currat, M. Ruedi, R. J. Petit, L. Excoffier, *Evolution* **62**, 1908 (2008).
84. J. Zilhão *et al.*, *PLoS ONE* **5**, e8880 (2010).
85. J. Krause *et al.*, *Curr. Biol.* **20**, 231 (2010).
86. P. Gunz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6094 (2009).
87. W. H. Li, C. I. Wu, C. C. Luo, *Mol. Biol. Evol.* **2**, 150 (1985).
88. We thank E. Buglione, A. Burke, Y.-J. Chen, J. Salem, P. Schaffer, E. Szekeres, and C. Turcotte at 454 Life Sciences Corp. for production sequencing on the 454 platform; S. Fisher, J. Wilkinson, J. Blye, R. Hegarty, A. Allen, S. K. Young, and J. L. Chang for nine Illumina sequencing runs performed at the Broad Institute; J. Rothberg and E. Rubin for input leading up to this project; O. Bar-Yosef, L. Excoffier, M. Gralle, J.-J. Hublin, D. Lieberman, M. Stoneking, and L. Vigilant for constructive criticism; I. Janković for assistance with the Vindija collection; S. Ptak, M. Siebauer, and J. Visagie for help with data analysis, M. Richards and S. Talamo for carbon dating; J. Dabney for editorial assistance; the Genome Center at Washington University for prepublication use of the orangutan genome assembly; and K. Finstermeier for expert graphical design. Neandertal bone extract sequence data have been deposited at European Bioinformatics Institute under STUDY accession ERP000119, alias Neandertal Genome project. HGDP sequence data have been deposited at EBI under STUDY accession ERP000121, alias Human Genome Diversity Project. We are grateful to the Max Planck Society, and particularly the Presidential Innovation Fund, for making this project possible. C.L.-F. was supported by a grant from the Ministerio de Ciencia e Innovación; E.Y.D. and M.S. were supported in part by grant GM40282; A.-S.M. was supported by a Janggen-Pöhn fellowship; N.F.H. and J.C.M. were supported in part by the Intramural Research Program of the National Human Genome Research Institute, National Institutes of Health; and D.R. by a Burroughs Wellcome Career Development Award in the Biomedical Sciences. Author contributions: S.P. conceived and coordinated the project; D.R. coordinated population genetic analyses; R.E.G. and J.K. coordinated bioinformatic aspects; R.E.G., J.Kr., A.W.B., M.E., and S.P. developed the initial project strategies; J.Kr. and T.M. collected and analyzed fossil samples; J.Kr., T.M., A.W.B., and M.M. developed the DNA extraction and library preparation protocols and performed laboratory work prior to sequencing; K.P. designed the restriction enzyme enrichment method; A.A.-P., A.B., B.Hb., B.Hf., M.Sg., R.S., A.W., J.A., M.E., and M.K. performed and coordinated DNA sequencing on the 454 and Illumina platforms; J.A. and M.E. organized and coordinated sequence production on the 454 platform; C.N., E.S.L., C.R., and N.N. organized and performed nine sequencing runs on the Illumina platform at the Broad Institute; M.K. and J.K. compiled the catalog of human-specific genomic features; U.S., M.K., N.H., J.M., J.K., K.P., and R.E.G. developed and implemented the primary sequence alignment and analysis methodologies; R.E.G., U.S., J.Kr., A.W.B., H.B., P.L.F.J. and M.L. developed and implemented the wet lab and bioinformatic assays for human DNA contamination; C.A., T.M.-B., and E.E.E. performed structural variation analyses; H.L., J.M., and D.R. designed and implemented analyses of population divergences; R.E.G., N.P., W.Z., J.M., H.L., M.H.-Y.F., E.Y.D., A.S.-M., P.L.F.J., J.J., J.G., M.L., D.F., M.S., E.B., R.N., S.P., and D.R. developed and implemented population genetics comparisons; R.E.G., M.L., J.G., D.F., J.D.J., D.R., and S.P. designed and implemented the screen for selective sweeps; P.R., D.B., Z.K., I.G., C.V., V.B.D., L.V.G., C.L.-F., M.R., J.F., A.R., and R.S. provided samples, analyses, and paleontological expertise; D.R. and S.P. edited the manuscript.

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/710/DC1

Materials and Methods

SOM Text

Figs. S1 to S51

Tables S1 to S58

References

8 February 2010; accepted 2 April 2010

10.1126/science.1188021

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/723>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/328/5979/723/DC1>

This article **cites 22 articles**, 9 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/328/5979/723#otherarticles>

This article has been **cited by** 1 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/cgi/content/full/328/5979/723#otherarticles>

This article appears in the following **subject collections**:

Genetics

<http://www.sciencemag.org/cgi/collection/genetics>

Targeted Investigation of the Neandertal Genome by Array-Based Sequence Capture

Hernán A. Burbano,^{1*} Emily Hodges,^{2,3*} Richard E. Green,^{1†} Adrian W. Briggs,¹ Johannes Krause,¹ Matthias Meyer,¹ Jeffrey M. Good,^{1,4} Tomislav Maricic,¹ Philip L. F. Johnson,⁵ Zhenyu Xuan,^{2‡} Michelle Rooks,^{2,3} Arindam Bhattacharjee,⁶ Leonardo Brizuela,⁶ Frank W. Albert,¹ Marco de la Rasilla,⁷ Javier Fortea,^{7§} Antonio Rosas,⁸ Michael Lachmann,¹ Gregory J. Hannon,^{2,3} Svante Pääbo¹

It is now possible to perform whole-genome shotgun sequencing as well as capture of specific genomic regions for extinct organisms. However, targeted resequencing of large parts of nuclear genomes has yet to be demonstrated for ancient DNA. Here we show that hybridization capture on microarrays can successfully recover more than a megabase of target regions from Neandertal DNA even in the presence of ~99.8% microbial DNA. Using this approach, we have sequenced ~14,000 protein-coding positions inferred to have changed on the human lineage since the last common ancestor shared with chimpanzees. By generating the sequence of one Neandertal and 50 present-day humans at these positions, we have identified 88 amino acid substitutions that have become fixed in humans since our divergence from the Neandertals.

The fossil record provides a rough chronological overview of the major phenotypic changes during human evolution. However, the underlying genetic bases for most of these events remain elusive. This is partly because it is not known when most human-specific genetic changes, identified from genome comparisons to living relatives, occurred during the ~6.5 million years since the separation of the human and chimpanzee evolutionary lineages. However, shotgun sequencing of the Neandertal, a human form whose ancestors split from modern human ancestors 270,000 to 440,000 years ago, has been performed to ~1.3-fold coverage of the entire genome (1). Comparison of Neandertal and present-day human genomes can reveal information about whether genetic changes occurred before or after the ancestral population split of modern humans and Neandertals. However, low-coverage whole-genome shotgun sequencing inevitably leaves a substantial proportion of the

genome uncovered. Although deeper shotgun sequencing of one or a few individuals may produce higher coverage across the whole genome, simple shotgun approaches cannot economically retrieve specific loci from multiple individuals, both due to the size of the mammalian genome per se and to the very high proportion (up to 99.9%) of microbial DNA in the vast majority of ancient tissue remains, with the exception of some instances of preservation in permafrost (2, 3). Primer extension capture can isolate specific DNA sequences from multiple Neandertal individuals (4). However, although

useful for capture of small target regions such as mitochondrial DNA (mtDNA) (4, 5), this method is unlikely to be scalable up to megabase target regions, ruling out experiments such as the retrieval of exomes, large chromosomal regions, or validation of sites of interest identified in the low-coverage shotgun genome data.

Because microarrays can carry hundreds of thousands of probes, we investigated the use of massively parallel hybridization capture on glass slide microarrays (6, 7) on Neandertal DNA at thousands of genomic positions where nucleotide substitutions changing amino acids (nonsynonymous substitutions) have occurred on the human lineage since its split from chimpanzees. For any substitution that is fixed, i.e., occurs in all present-day humans, it is currently impossible to judge how long ago either the original mutation or the subsequent fixation event occurred. However, by ascertaining the Neandertal state at these positions, we can separate fixed substitutions into two classes: (i) sites where a Neandertal carries the derived state, which indicates that the substitution must have occurred before the population split of modern humans and Neandertals; and (ii) sites where a Neandertal is ancestral, which indicates that fixation of a substitution in modern humans occurred after the population split with Neandertals (Fig. 1A).

To identify substitutions that occurred on the human lineage since the ancestral split with chimpanzee, we aligned human, chimpanzee, and orangutan protein sequence for all orthologous proteins in HomoloGene (8, 9). Comparison of these three species allowed us to assign human/chimpanzee differences to their respective evolutionary lineages. We designed a 1 Million Agilent oligonucleotide array covering, at 3-base pair tiling, all 13,841 nonsynonymous substitutions

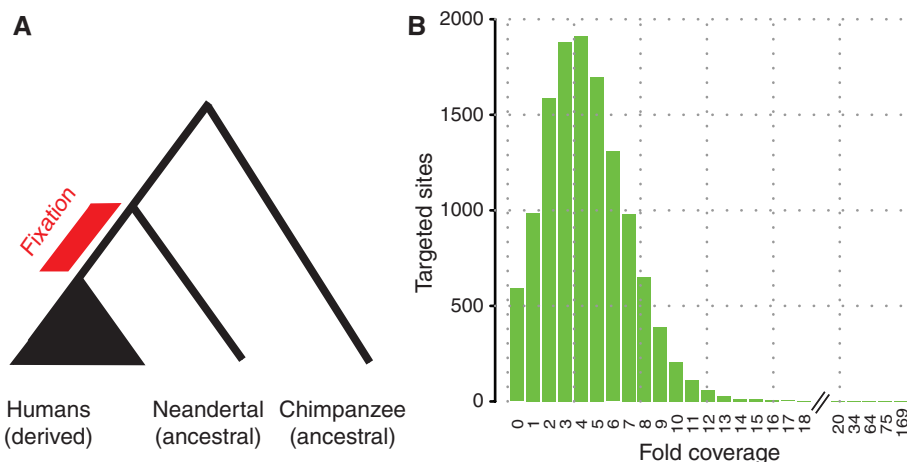


Fig. 1. (A) Identification of protein-coding changes that are likely to have become fixed recently (red bar) in modern humans after the population split from Neandertals. Such positions would be derived in all present-day humans but ancestral in the Neandertal. (B) Distribution of Neandertal coverage for ~14,000 amino acid substitution sites found in the human genome by comparison to primate outgroups. The same sites were also sequenced in 50 present-day humans. Of these, 88 were found to be fixed derived in present-day humans and ancestral in Neandertal, representing recently fixed protein-coding changes in the human genome.

¹Max Planck Institute for Evolutionary Anthropology, D-04103 Leipzig, Germany. ²Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. ³Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. ⁴Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA. ⁵Department of Biology, Emory University, Atlanta, GA 30322, USA. ⁶Agilent Technologies, Life Sciences Group, Santa Clara, CA 95051, USA. ⁷Área de Prehistoria, Departamento de Historia, Universidad de Oviedo, Oviedo, Spain. ⁸Departamento de Paleobiología, Museo Nacional de Ciencias Naturales, Consejo Superior de Investigaciones Científicas, Madrid, Spain.

*These authors contributed equally to this work.

†Present address: Department of Biomolecular Engineering, University of California, Santa Cruz, Santa Cruz, CA 95064, USA.

‡Present address: Department of Molecular and Cell Biology, Center for Systems Biology, University of Texas at Dallas, Richardson, TX 75080, USA.

§Deceased.

inferred to have occurred on the human lineage (9). We used this array to capture DNA from a ~49,000-year-old Neandertal bone (Sidrón 1253) from El Sidrón Cave, Spain (10, 11). This bone contains a high amount of Neandertal DNA in absolute terms, but also a high proportion (99.8%) of microbial DNA (4), making it unsuitable for shotgun sequencing. To identify which of the 13,841 substitutions are fixed in present-day humans, we also collected data from 50 individuals from the Human Genome Diversity Panel (12) with the same array design as used for the El Sidrón Neandertal (table S1). The DNA libraries from these individuals were barcoded, pooled, and captured on a single array (13). All captured products were sequenced on the Illumina GAII platform and aligned to the human genome (9). Overall, 37% of the Neandertal sequence reads aligned to the target regions, representing ~190,000-fold target enrichment. We retrieved Neandertal sequence for 13,250 (96%) of the substitutions targeted on the array, with an average coverage of 4.8-fold after filtering for polymerase chain reaction (PCR) duplicates (Fig. 1B). We considered a Neandertal position ancestral if all overlapping reads matched the chimpanzee state and derived if all reads carried the modern human state or if we found a mixture of derived and third-state reads, disregarding positions that carried only a third state or positions where Neandertal reads were found both in the ancestral and in the derived state. From each present-day individual, a total of 25% (23 to 27%) of reads aligned to the target regions. In each individual, we retrieved on average 98% (97 to 99%) of targeted positions and had on average coverage of 10-fold (fig. S1). We estimated genotypes for each individual and considered a position to be fixed derived if it was homozygous and derived in all humans ob-

served, and if data were available for at least 25 individuals (50 chromosomes) (9).

We included several additional target regions on the array to assess levels of human DNA contamination, which can frequently affect ancient DNA experiments (14). One such region was the complete human mtDNA, which is known to differ between the Sidrón 1253 Neandertal analyzed here and almost all (99%) present-day humans at 130 positions (4). Even though the array probes were designed to match present-day human mtDNA, 253,549 of the 254,296 (99.71%) fragments that overlapped these 130 positions matched the Neandertal state. We therefore conclude that the vast majority of mtDNA in the Sidrón 1253 library is of Neandertal origin.

For a more direct estimate of contamination in the nuclear DNA, we used 46 nucleotide sites on the X chromosome that differ between present-day humans and chimpanzees and that were found to be ancestral in a Neandertal from Croatia (Vindija 33.16) by shotgun sequencing (1), whereas ~1000 present-day humans in the human diversity panel carry a derived state. The Sidrón 1253 individual will obviously not match Vindija 33.16 at all of these sites. However, because Sidrón 1253 is a male (15) and thus carried a single X chromosome, at sites where he does match Vindija 33.16, all reads should carry the ancestral base while apparent heterozygosity will indicate human DNA contamination. By analyzing the consistency of reads overlapping these sites on the X chromosome, we calculated a maximum likelihood estimator of X-chromosomal contamination of 4%, although confidence intervals are large (1 to 12%) due to the small number of relevant positions (9).

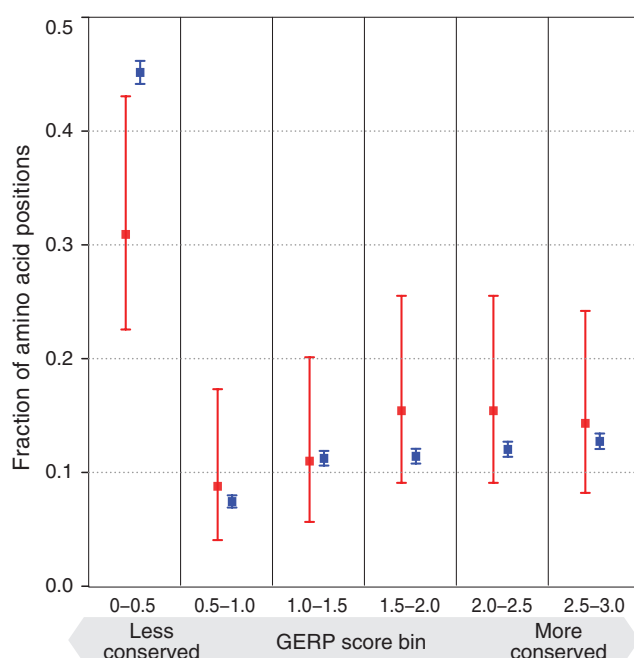
Another way to estimate contamination across autosomes is to investigate patterns of allele

counts. Because at every site an individual is either homozygous derived, homozygous ancestral, or heterozygous, DNA from a single individual will yield at each site either only derived alleles, only ancestral alleles, or a draw with equal chance for either. Contamination from other individuals would cause systematic deviation from these patterns. We thus produced a likelihood model that estimated contamination at the positions recovered from Sidrón 1253, and calculated a 95% upper bound for contamination of 2% (9). From these results we conclude that the Sidrón 1253 data are not substantially affected by human DNA contamination.

In total, we determined with high confidence the Neandertal and present-day human state for 10,952 nonsynonymous substitutions. In 10,015 (91.5%) of all cases the Neandertal carries the derived state, whereas in 937 (8.5%) cases the ancestral state was found (fig. S2). Of the positions that are fixed in the derived state in present-day humans, 9525 (87%) are derived in Neandertal, whereas 88 (0.8%) (table S2) are ancestral (fig. S2). In agreement with previous results generated by PCR (15), two substitutions that change amino acids in the gene *FOXP2* (16), involved in speech and language (17), are both derived in this Neandertal individual.

The 88 recently fixed substitutions occur in 83 genes (tables S2 and S3). We asked if these genes cluster in any group of functionally related genes relative to the genes that were targeted in the capture array (18) (as defined in the Gene Ontology) but found no such groups. We furthermore asked if the 88 substitutions that recently became fixed in humans differ from those that occurred before the divergence from the Neandertal with respect to how evolutionarily conserved the positions in the encoded proteins are (9, 19) (Fig. 2). We found that the 88 recent substitutions tend to affect amino acid positions that are more conserved than the older substitutions (Wilcoxon rank test; $P = 0.014$). Similarly, the recently fixed substitutions caused more radical amino acid changes with respect to the chemical properties of the amino acids (Wilcoxon rank test; $P = 0.04$). One possible explanation for these observations is that the effective population size of humans since their separation from the Neandertal lineage has been small, leading to a reduced efficiency of purifying selection, as seen, e.g., in Europeans (20). We also looked for evidence that the recent substitutions may have been fixed by positive selection. One recent substitution occurred in *SCML1*, a gene involved in spermatogenesis (21) that has been previously proposed as a target of positive selection in humans (22) as well as frequent positive selection in primates (23). However, we found no significant overrepresentation of the 83 genes among candidate genes in three genome-wide scans for positive selection (24) (table S4). Nevertheless, we believe that all of these amino acid substitutions warrant functional studies.

Fig. 2. Evolutionary conservation at positions affected by substitutions that are fixed in present-day humans. For each bin of conservation GERP (Genomic Evolutionary Rate Profiling) scores, the fractions of derived and ancestral alleles of all positions where the Neandertal carries derived (blue) and ancestral alleles (red), respectively, are given. Error bars are 95% binomial confidence intervals.



Our results demonstrate that hybridization capture arrays can generate data from genomic target regions of megabase size from ancient DNA samples, even when only ~0.2% of the DNA in a sample stems from the endogenous genome. By generating an average coverage of 4- to 5-fold, errors from sequencing and small amounts of human DNA contamination can be minimized. A further approximately 5-fold reduction of errors was achieved here by the enzymatic removal of uracil residues that are frequent in ancient DNA (25). Because the Sidrón 1253 Neandertal library used for this study has been amplified and effectively immortalized, the same library should be able to provide similar-quality data for any other genomic target region, or even the entire single-copy fraction of the Neandertal genome.

References and Notes

1. R. E. Green *et al.*, *Science* **328**, 710 (2010).
2. H. N. Poinar *et al.*, *Science* **311**, 392 (2006).
3. W. Miller *et al.*, *Nature* **456**, 387 (2008).
4. A. W. Briggs *et al.*, *Science* **325**, 318 (2009).
5. J. Krause *et al.*, *Curr. Biol.* **20**, 231 (2010).
6. E. Hodges *et al.*, *Nat. Protoc.* **4**, 960 (2009).
7. E. Hodges *et al.*, *Nat. Genet.* **39**, 1522 (2007).
8. E. W. Sayers *et al.*, *Nucleic Acids Res.* **38** (Database issue), D5 (2010).
9. Materials and methods are available as supporting material on Science Online.
10. A. Rosas *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19266 (2006).
11. T. De Torres *et al.*, *Archaeometry* **10.1111/j.1475-4754.2009.00491.x** (2009).
12. H. M. Cann *et al.*, *Science* **296**, 261 (2002).
13. M. Meyer, M. Kircher, *Cold Spring Harb. Protoc.* **10.1101/pdb.prot5448** (2010).
14. R. E. Green *et al.*, *EMBO J.* **28**, 2494 (2009).
15. J. Krause *et al.*, *Curr. Biol.* **17**, 1908 (2007).
16. W. Enard *et al.*, *Nature* **418**, 869 (2002).
17. F. Vargha-Khadem, D. G. Gadian, A. Copp, M. Mishkin, *Nat. Rev. Neurosci.* **6**, 131 (2005).
18. K. Prüfer *et al.*, *BMC Bioinformatics* **8**, 41 (2007).
19. G. M. Cooper *et al.*, *Genome Res.* **15**, 901 (2005).
20. K. E. Lohmueller *et al.*, *Nature* **451**, 994 (2008).
21. B. Boeckmann *et al.*, *Nucleic Acids Res.* **31**, 365 (2003).
22. C. D. Bustamante *et al.*, *Nature* **437**, 1153 (2005).
23. H. H. Wu, B. Su, *BMC Evol. Biol.* **8**, 192 (2008).
24. J. M. Akey, *Genome Res.* **19**, 711 (2009).
25. A. W. Briggs *et al.*, *Nucleic Acids Res.* **38**, e87 (2010).
26. We thank C. S. Burbano, C. de Filippo, J. Kelso, and D. Reich for helpful comments; M. Kircher, K. Prüfer, and U. Stenzel for technical support; C. D. Bustamante and K. E. Lohmueller for access to human resequencing databases; D. L. Goode and A. Sidow for providing conservation scores; J. M. Akey for providing coordinates of genome-wide scans for selection; I. Gut for human genotyping; E. Leproust and M. Srinivasan for providing early access to the 1 Million feature Agilent microarrays; and the Genome Center at Washington University for pre-publication use of the orangutan genome assembly (http://genome.wustl.edu/genomes/view/pongo_abelii/). The government of the Principado de Asturias funded excavations at the Sidrón site. J.M.G. was supported by an NSF international postdoctoral fellowship (OISE-0754461) and E.H. by a postdoctoral training grant from the NIH and by a gift from the Stanley Foundation. G.J.H. is an investigator of the Howard Hughes Medical Institute, which together with the Presidential Innovation Fund of the Max Planck Society provided generous financial support. DNA sequences are deposited in the European Bioinformatics Institute short read archive, with accession number ERP000125. The array capture technologies used in this study are the subject of pending patent filings U.S. 60/478, 382 (filed 2003) and U.S. 61/205, 834 (filed 2009), on which G.J.H. and E.H. are listed as inventors.

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/723/DC1
Materials and Methods
Figs. S1 to S4
Tables S1 to S5

8 February 2010; accepted 1 April 2010
10.1126/science.1188046

Fermi Gamma-Ray Imaging of a Radio Galaxy

The Fermi-LAT Collaboration*

The Fermi Gamma-ray Space Telescope has detected the γ -ray glow emanating from the giant radio lobes of the radio galaxy Centaurus A. The resolved γ -ray image shows the lobes clearly separated from the central active source. In contrast to all other active galaxies detected so far in high-energy γ -rays, the lobe flux constitutes a considerable portion (greater than one-half) of the total source emission. The γ -ray emission from the lobes is interpreted as inverse Compton-scattered relic radiation from the cosmic microwave background, with additional contribution at higher energies from the infrared-to-optical extragalactic background light. These measurements provide γ -ray constraints on the magnetic field and particle energy content in radio galaxy lobes, as well as a promising method to probe the cosmic relic photon fields.

Centaurus A (Cen A) is one of the brightest radio sources in the sky and was among the first identified with a galaxy (NGC 5128) outside of our Milky Way (1). Straddling the bright central source is a pair of extended radio lobes with a total angular extent of $\sim 10^\circ$ (2, 3), which makes Cen A the largest discrete nonthermal extragalactic radio source visible from Earth. At a distance of 3.7 Mpc (4), it is the nearest radio galaxy to Earth, and the implied physical source size is ~ 600 kpc. Such double-lobed radio structures associated with otherwise apparently normal giant elliptical galaxies have become the defining feature of radio galaxies in general. The

consensus explanation for this phenomenon is that the lobes are fueled by relativistic jets produced by accretion activity in a super-massive black hole residing at the galaxy's center.

With its unprecedented sensitivity and imaging capability (per-photon resolution: $\theta_{68} \approx 0.8 E_{\text{GeV}}^{-0.8}$), the Fermi Large Area Telescope (LAT) (5) has detected and imaged the radio lobes of Cen A in high-energy γ -rays. The LAT image resulting from ~ 10 months of all-sky survey data (Fig. 1) clearly shows the γ -ray peak coincident with the active galactic nucleus detected by the Compton/EGRET instrument (6) and extended emission from the southern giant lobe. Because the northern lobe is characterized by lower surface-brightness emission (in radio), it is not immediately apparent from a naked-eye inspection of the γ -ray counts map. Nevertheless, from a counts profile extracted along the north-south axis of the source (Fig. 2), γ -ray excesses from both lobes are clearly visible.

Spectra for each of the lobes together with the central source (hereafter referred to as the "core") were determined with a binned maximum likelihood analysis implemented in GTLIKE (7) using events from 0.2 to 30 GeV in equal logarithmically spaced energy bins. We modeled background emission by including the Galactic diffuse component, an isotropic component, and nearby γ -ray point sources [see the supporting online material (SOM)]. We fit the core as a point source at the known radio position and modeled the lobe emission with a 22-GHz Wilkinson Microwave Anisotropy Probe (WMAP) image (Fig. 1) (8) with the core region within a 1° radius excluded as a spatial template. The modeled lobe region roughly corresponds to the regions 1 and 2 (north) and 4 and 5 (south) defined in (9), where region 3 is the core (Fig. 2). Assuming a power law for the γ -ray spectra, we find a large fraction ($>1/2$) of the total >100 -MeV emission from Cen A to originate from the lobes with the flux in each of the northern $\{[0.77(+0.23/-0.19)_{\text{stat.}}(+0.39)_{\text{syst.}}] \times 10^{-7} \text{ ph cm}^{-2} \text{ s}^{-1}\}$ and southern $\{[1.09(+0.24/-0.21)_{\text{stat.}}(+0.32)_{\text{syst.}}] \times 10^{-7} \text{ ph cm}^{-2} \text{ s}^{-1}\}$ lobes smaller than the core flux $\{[1.50(+0.25/-0.22)_{\text{stat.}}(+0.37)_{\text{syst.}}] \times 10^{-7} \text{ ph cm}^{-2} \text{ s}^{-1}\}$ (stat., statistical; syst., systematic). Uncertainties in the LAT effective area, the Galactic diffuse model used, and the core exclusion region were considered to be sources of systematic error (SOM). The resultant test statistic (10) for the northern and southern giant lobes are 29 and 69, which correspond to detection significances of 5.0σ and 8.0σ , respectively. The lobe spectra are steep, with photon indices $\Gamma = 2.52(+0.16/-0.19)_{\text{stat.}}(+0.25)_{\text{syst.}}$ (north) and $2.60(+0.14/-0.15)_{\text{stat.}}(+0.20)_{\text{syst.}}$ (south) in which photons up to ~ 2 to 3 GeV are currently detected. These values are consistent with that of the core [$\Gamma = 2.67(+0.10)_{\text{stat.}}(+0.08)_{\text{syst.}}$], which

*All authors with their affiliations appear at the end of this paper.

†To whom correspondence should be addressed. E-mail: Teddy.Cheung.ctr@nrl.navy.mil (C.C.C.); fukazawa@hep01.hepl.hiroshima-u.ac.jp (Y.F.); jurgen.knodlseder@cesr.fr (J.K.); stawarz@slac.stanford.edu (Ł.S.)

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/725>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/science.1184656/DC1>

This article appears in the following **subject collections**:

Astronomy

<http://www.sciencemag.org/cgi/collection/astronomy>

Fig. 1. (A and B) Fermi-LAT γ -ray (>200 MeV) counts maps centered on Cen A, displayed with square-root scaling. In both (A) and (B), models of the galactic and isotropic emission components were subtracted from the data (in contrast to the observed counts profile presented in Fig. 2). The images are shown before (A) and after (B) additional subtraction of field point sources (SOM) and are shown adaptively smoothed with a minimum signal-to-noise ratio of 10. In (B), the white circle with a diameter of 1° is approximately the scale of the LAT point-spread function width. (C) For comparison, the 22-GHz radio map from the 5-year WMAP data set (8) with a resolution of $0^\circ.83$ is shown. J2000, equinox; h, hour; m, minutes.

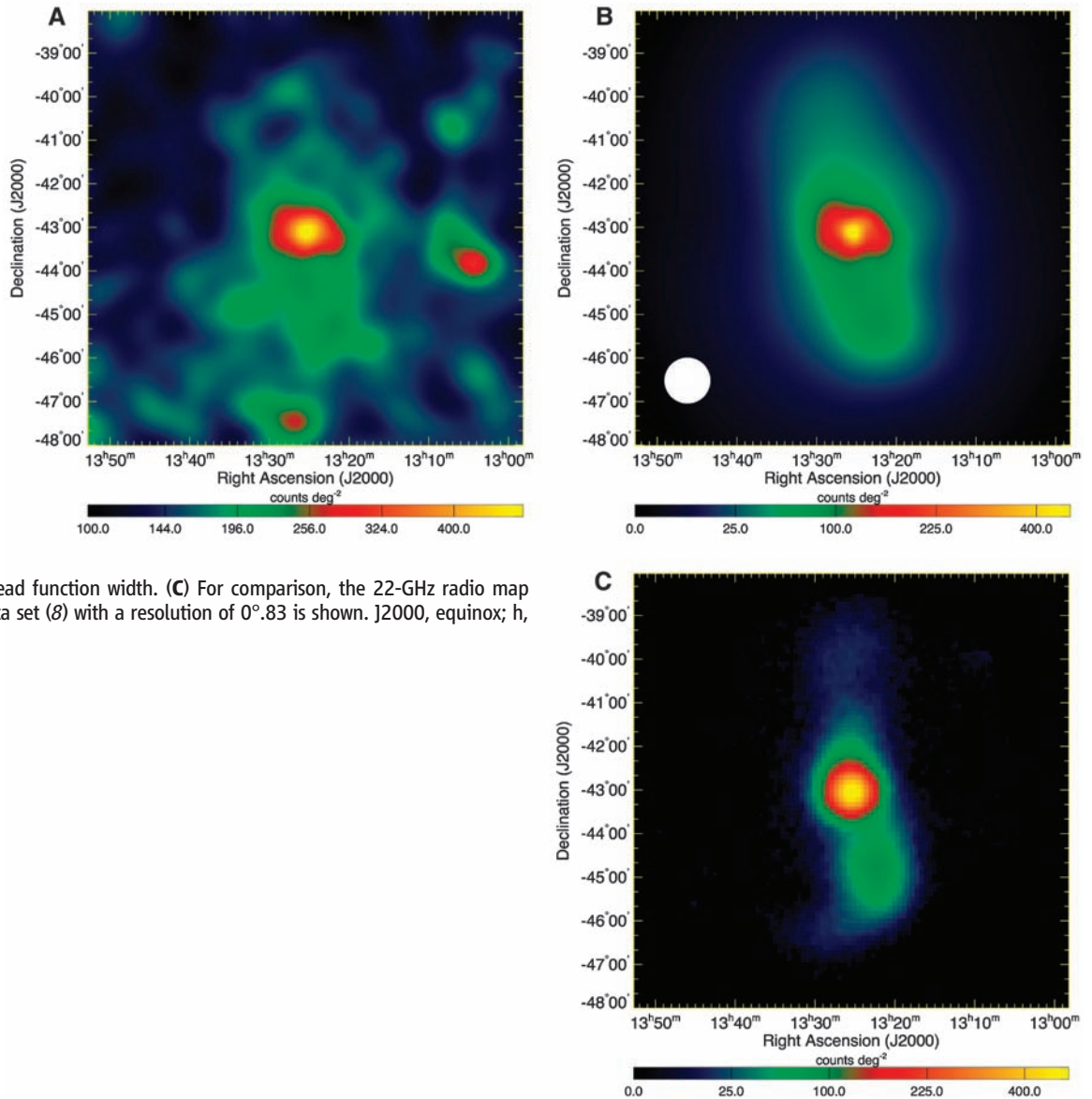
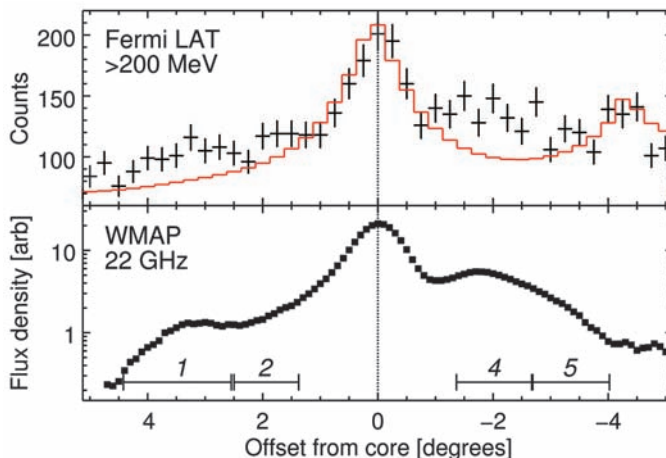


Fig. 2. Observed intensity profiles of Cen A along the north-south axis in γ -rays (top) and in the radio band (bottom). In the bottom panel, the lobe regions 1 and 2 (northern lobe) and regions 4 and 5 (southern lobe) are indicated as in (9), where region 3 (not displayed here) is the core. The red curve overlaid onto the LAT data indicates the emission model for all fitted point sources, plus the isotropic and Galactic diffuse (brighter to the south) emission. The point sources include the Cen A core (offset = 0°) and a LAT source (offset = -4.5°) (see SOM) that is clearly outside (1° from the southern edge) of the southern lobe. The excess counts are coincident with the northern and southern giant lobes. arb, arbitrary units.



is known to have a steep γ -ray spectrum (6). For further details pertaining to the analysis of the lobe emission, see the SOM.

It is well-established that radio galaxy lobes are filled with magnetized plasma containing ultra-relativistic electrons emitting synchrotron radiation in the radio band (observed frequencies: $\nu \sim 10^7$ to 10^{11} Hz). These electrons also up-scatter ambient photons to higher energies via the inverse Compton (IC) process. At the observed distances far from the parent galaxy (>100 -kpc scale), the dominant soft-photon field surrounding the extended lobes is the pervading radiation from the cosmic microwave background (CMB) (11). Because IC/CMB scattered emission in the lobes of more distant radio galaxies is generally well observed in the x-ray band (12–14), the IC spectrum can be expected to extend to even higher energies (9, 15), as demonstrated by the LAT detection of the Cen A giant lobes.

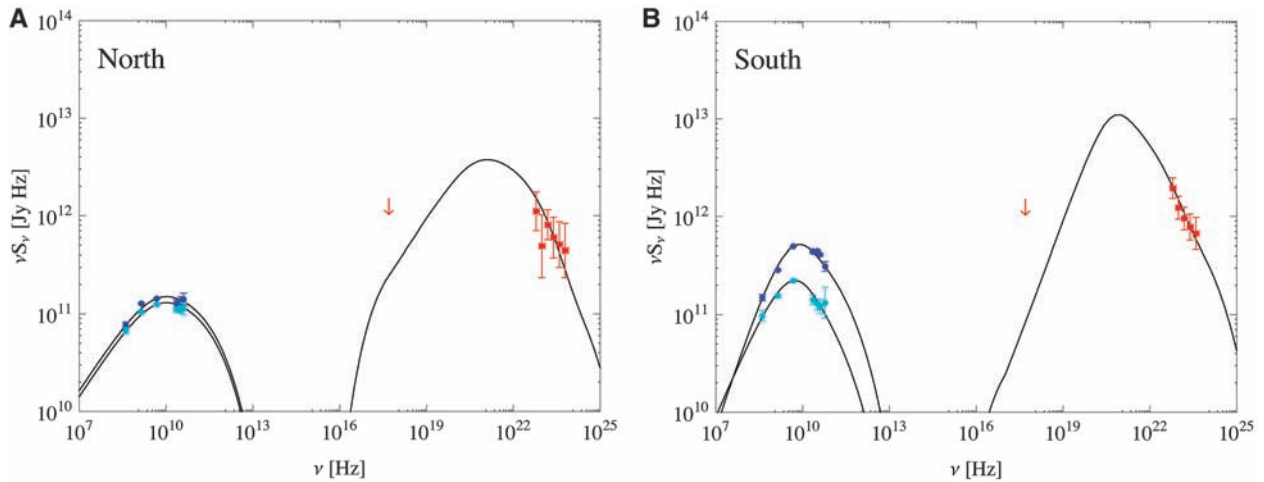


Fig. 3. Broad-band SEDs of the northern (**A**) and southern (**B**) giant lobes of Cen A. The radio measurements (up to 60 GHz) of each lobe are separated into two regions, with dark blue data points indicating regions that are closer to the nucleus (regions 2 and 4; see Fig. 2), and light blue points denoting the farther regions (1 and 5). Synchrotron continuum models for each region are overlaid. The component at higher energies is the total IC

emission of each lobe modeled to match the LAT measurements (red points with error bars; error bars indicate 1σ errors). The x-ray limit for the lobe emission derived from *SAS-3* observations (24) is indicated with a red arrow [see (9)]. The break and maximum frequencies in the synchrotron spectra are $\nu_{br} = 4.8$ GHz and $\nu_{max} = 400$ GHz, respectively. νS_ν , frequency multiplied by flux density.

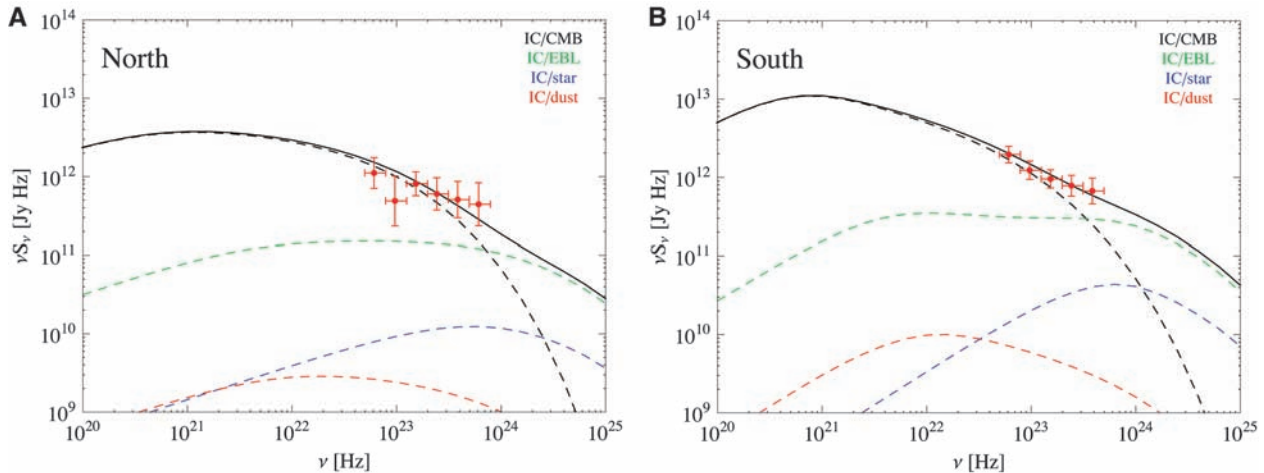


Fig. 4. Detail of the IC portion of the northern (**A**) and southern (**B**) giant lobes' SEDs (Fig. 3). The separate contributions from the different photon seed sources are indicated with dashed lines, and the total emission is

represented by the solid black line. Red data points and error bars are the same as in Fig. 3. Vertical bars indicate errors; horizontal bars indicate frequency range.

To model the observed lobe γ -rays as IC emission, detailed radio measurements of the lobes' synchrotron continuum spectra are necessary to infer the underlying electron energy distribution (EED), $n_e(\gamma)$, where the electron energy is $E_e = \gamma m_e c^2$ (γ , electron Lorentz factor; m_e , electron mass; c , speed of light; n_e , number density of electrons). In anticipation of these Fermi observations, ground-based (16, 17) and WMAP satellite (8) maps of Cen A were previously analyzed (9). Here, we separately fit the 0.4- to 60-GHz measurements for each region defined therein for the north (1 and 2) and south (4 and 5) lobes (Fig. 2) with EEDs in the form of a broken power law (with normalization k_e and slopes s_1 and s_2) plus an exponential cutoff at high energies $n_e(\gamma) = k_e \gamma^{-s_1}$ for $\gamma_{min} \leq \gamma < \gamma_{br}$ and $n_e(\gamma) = k_e \gamma_{br}^{s_2-s_1} \gamma^{-s_2} \exp[-\gamma/\gamma_{max}]$ for $\gamma \geq \gamma_{br}$, such that the electron energy density is $U_e = \int E_e n_e(\gamma) d\gamma$. To a certain extent, our modeling results depend on the shape of the electron spectrum

at energies higher than those probed by the WMAP measurements ($\nu \geq 60$ GHz) (Fig. 3); we have assumed the spectrum to decline exponentially.

We calculated the IC spectra resulting from the fitted EED (parameters listed in table S1 of the SOM) by employing precise synchrotron (18) and IC (19) kernels (including Klein-Nishina effects) by adjusting the magnetic field B . In addition to the CMB photons, we included IC emission off the isotropic infrared-to-optical extragalactic background light (EBL) radiation field (9, 20, 21), using the data compilation from (22). Anisotropic radiation from the host galaxy starlight and the well-known dust lane was also included, but was found to have a negligible contribution in comparison to the EBL (Fig. 4 and SOM). The resultant total IC spectra of the northern and southern lobes (Fig. 3) with $B = 0.89$ μ G (north) and 0.85 μ G (south) provide satisfactory representations of

the observed γ -ray data. These B -field values imply that the high-energy γ -ray emission detected by the LAT is dominated by the scattered CMB emission, with the EBL contributing at higher energies (≥ 1 GeV) (Fig. 4).

Considering only contributions from ultra-relativistic electrons and magnetic field, the lobe plasma is found to be close to the minimum-energy condition with the ratio of the energy densities $U_e/U_B \approx 4.3$ (north) and ≈ 1.8 (south), where $U_B = B^2/8\pi$. The EED was assumed to extend down to $\gamma_{min} = 1$; adopting larger values can reduce this ratio by a fractional amount for the southern lobe and by up to \sim two times for the northern lobe (SOM). For comparison, IC/CMB x-ray measurements of extended lobes of more powerful [Fanaroff-Riley type-II (23)] radio sources have been used to infer higher B fields and equipartition ratios with a range $U_e/U_B \approx 1$ –10 (12–14).

The radiating particles in the Cen A lobes lose energy predominantly through the IC channel, because the ratio of the corresponding cooling times is equal to the energy density ratio $U_{\text{CMB}}/U_B \gtrsim 10$. This manifests itself in the approximately one order of magnitude dominance of the γ -ray component over the radio component in the observed spectral energy distributions (SEDs) (Fig. 3). However, the magnetic-field constraints (thus, the exact ratios of U_{CMB}/U_B) are sensitive to the shape of the EED at the electron energies $E_e > 0.1$ TeV. On one hand, magnetic-field strengths greater than $B \sim 1$ μG will under-produce the observed LAT emission for all reasonable forms of the EED, so the quoted ratio is formally a lower limit. Conversely, magnetic fields as low as \sim one-third of our quoted values are strictly allowed if we invoke a sharper cutoff in the synchrotron spectrum at $\gtrsim 60$ GHz, as would be expected in some aging models for extended radio lobes (9). Such models with lower magnetic fields and EEDs with sharper upper-energy cutoffs than the exponential form adopted here (Fig. 3) would result in IC spectra in which the EBL, rather than the CMB, component becomes dominant in the LAT observing band. These models require large departures from equipartition ($U_e/U_B \gtrsim 10$); even lower B fields would violate the observed x-ray limit to the lobe flux (9, 24).

For a tangled magnetic field, the total nonthermal pressures in the lobes are $p_{\text{nt}} = (U_e + U_B)/3 \approx 5.6 \times 10^{-14}$ erg cm $^{-3}$ (north) and $\approx 2.7 \times 10^{-14}$ erg cm $^{-3}$ (south). Such estimates can be compared to the ambient thermal gas pressure to enable further understanding of the dynamical evolution of such giant structures in general. Unfortunately, the parameters of the thermal gas at the appropriate distances from the nucleus of Cen A are not well known. Upper limits of the soft x-ray emission of the lobes (9), as well as Faraday rotation studies (25), indicate that the thermal gas number density is $n_{\text{gas}} < 10^{-4}$ cm $^{-3}$ within the giant lobes. Hence, the upper limit for the thermal pressure $p_{\text{gas}} = nkT < 10^{-13}(n_{\text{gas}}/10^{-4} \text{ cm}^{-3})(T_{\text{gas}}/10^7 \text{ K})$ erg cm $^{-3}$ (k , the Boltzmann constant; T , temperature) is comparable to the evaluated nonthermal pressures.

Our modeling results allow us to estimate the total energy in both giant lobes $E_{\text{tot}} 1.5 \times 10^{58}$ erg. This energy, divided by the lifetime of the lobes derived from spectral aging $\tau \approx 3 \times 10^7$ years (9), gives the required kinetic power of the jets inflating the giant lobes, $L_j \approx E_{\text{tot}}/2\tau \approx 7.7 \times 10^{42}$ erg s $^{-1}$, which is close to the estimates of the total power of the kiloparsec-scale outflow in the current epoch of jet activity (26). For a black hole mass in Cen A, $M_{\text{BH}} \approx 10^8 M_{\odot}$ (M_{\odot} , mass of the sun) (27), this implies a jet power that is only a small fraction of the Eddington luminosity ($L_j \approx 6.1 \times 10^{-4} L_{\text{Edd}}$), as well as a relatively small jet production efficiency ($E_{\text{tot}}/M_{\text{BH}}c^2 \approx 8 \times 10^{-5}$). Because the work done by the expanding lobes on the ambient medium is not taken into account and the

relativistic proton content is unconstrained in our analysis, the obtained values for E_{tot} and L_j are strict lower limits and could plausibly be an order of magnitude larger (28).

The observed LAT emission implies the presence of 0.1 to 1 TeV electrons in the few hundred kiloparsec-scale lobes. Because their radiative lifetimes (< 1 to 10 million years) approach plausible electron transport time scales across the lobes, the particles have been either accelerated in situ or efficiently transported from regions closer to the nucleus. Such high-energy electrons in the lobes are, in fact, required to IC scatter photons into the LAT band, and it is presently unclear how common this is in other radio galaxies.

References and Notes

1. J. G. Bolton, G. J. Stanley, O. B. Slee, *Nature* **164**, 101 (1949).
2. C. A. Shain, *Aust. J. Phys.* **11**, 517 (1958).
3. K. V. Sheridan, *Aust. J. Phys.* **11**, 400 (1958).
4. L. Ferrarese et al., *Astrophys. J.* **654**, 186 (2007).
5. W. B. Atwood et al., *Astrophys. J.* **697**, 1071 (2009).
6. R. C. Hartman et al., *Astrophys. J. Suppl. Ser.* **123**, 79 (1999).
7. <http://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Cicerone/>
8. G. Hinshaw et al., *Astrophys. J. Suppl. Ser.* **180**, 225 (2009).
9. M. J. Hardcastle, C. C. Cheung, I. J. Feain, L. Stawarz, *Mon. Not. R. Astron. Soc.* **393**, 1041 (2009).
10. J. R. Mattox et al., *Astrophys. J.* **461**, 396 (1996).
11. D. E. Harris, J. E. Grindlay, *Mon. Not. R. Astron. Soc.* **188**, 25 (1979).
12. E. D. Feigelson, S. A. Laurent-Muehleisen, R. I. Kollgaard, E. B. Fomalont, *Astrophys. J.* **449**, L149 (1995).
13. J. H. Croston et al., *Astrophys. J.* **626**, 733 (2005).
14. J. Kataoka, L. Stawarz, *Astrophys. J.* **622**, 797 (2005).
15. C. C. Cheung, *AIP Conf. Ser.* **921**, 325 (2007).
16. N. Junkes, R. F. Haynes, J. I. Harnett, D. L. Jauncey, *Astron. Astrophys.* **269**, 29 (1993).
17. H. Alvarez, J. Aparici, J. May, P. Reich, *Astron. Astrophys.* **355**, 863 (2000).
18. A. Crusius, R. Schlickeiser, *Astron. Astrophys.* **164**, L16 (1986).
19. G. R. Blumenthal, R. J. Gould, *Rev. Mod. Phys.* **42**, 237 (1970).
20. M. G. Hauser, E. Dwek, *Annu. Rev. Astron. Astrophys.* **39**, 249 (2001).
21. M. Georganopoulos et al., *Astrophys. J.* **686**, L5 (2008).
22. M. Raue, D. Mazin, *Int. J. Mod. Phys. D* **17**, 1515 (2008).
23. B. L. Fanaroff, J. M. Riley, *Mon. Not. R. Astron. Soc.* **167**, 31P (1974).
24. F. J. Marshall, G. W. Clark, *Astrophys. J.* **245**, 840 (1981).
25. I. J. Feain et al., *Astrophys. J.* **707**, 114 (2009).
26. J. H. Croston et al., *Mon. Not. R. Astron. Soc.* **395**, 1999 (2009).
27. A. Marconi et al., *Astron. Astrophys.* **448**, 921 (2006).
28. C. D. Dermer, S. Razzaque, J. D. Finke, A. Atoyan, *N. J. Phys.* **11**, 065016 (2009).
29. The Fermi-LAT Collaboration acknowledges support from a number of agencies and institutes for both the development and operation of the LAT, as well as for scientific data analysis. These organizations include NASA and the U.S. Department of Energy in the United States; Commissariat à l'Énergie Atomique and the Centre National de la Recherche Scientifique/Institut National de Physique Nucléaire et de Physique des Particules in France; the Agenzia Spaziale Italiana and the Istituto Nazionale di Fisica Nucleare in Italy; the Ministry of Education, Culture, Sports, Science and Technology (MEXT), High Energy Accelerator Research Organization (KEK), and Japan Aerospace Exploration

Agency (JAXA) in Japan; and the K. A. Wallenberg Foundation, the Swedish Research Council, and the National Space Board in Sweden. Additional support from Istituto Nazionale di Astrofisica in Italy and Centre National d'Études Spatiales in France for science analysis during the operations phase is also gratefully acknowledged. C.C.C. was supported by the NASA Postdoctoral Program at Goddard Space Flight Center, administered by Oak Ridge Associated Universities through a contract with NASA. J. Conrad is a Royal Swedish Academy of Sciences research fellow (funded by a grant from the K. A. Wallenberg Foundation). We thank N. Odegard for providing the WMAP image.

The Fermi-LAT Collaboration

A. A. Abdo,^{1,2} M. Ackermann,³ M. Ajello,³ W. B. Atwood,⁴ L. Baldini,⁵ J. Ballet,⁶ G. Barbiellini,^{7,8} D. Bastieri,^{9,10} B. M. Baughman,¹¹ K. Bechtol,⁵ R. Bellazzini,⁵ B. Berenji,³ R. D. Blandford,³ E. D. Bloom,³ E. Bonamente,^{12,13} A. W. Borgland,³ J. Bregeon,³ A. Brez,⁵ M. Brigida,^{14,15} P. Bruel,¹⁶ T. H. Burnett,¹⁷ S. Buson,¹⁰ G. A. Caliendo,¹⁸ R. A. Cameron,³ P. A. Caraveo,¹⁹ J. M. Casandjian,⁶ E. Cavazzuti,²⁰ C. Cecchi,^{12,13} Ö. Celik,^{21,22,23} A. Chekhtman,^{1,24} C. C. Cheung,^{1,22,21} J. Chiang,³ S. Ciprini,¹³ R. Claus,⁵ J. Cohen-Tanugi,²⁵ S. Colafrancesco,²⁰ L. R. Cominsky,²⁶ J. Conrad,^{27,28} L. Costamante,³ S. Cutini,²⁰ S. D. Davis,^{21,23} C. D. Dermer,¹ A. de Angelis,²⁹ F. de Palma,^{14,15} S. W. Digel,³ E. do Couto e Silva,³ P. S. Drell,³ R. Dubois,³ D. Dumora,^{30,31} C. Farnier,²⁵ C. Favuzzi,^{14,15} S. J. Fegan,¹⁶ J. Finke,^{1,2} W. B. Focke,³ P. Fortin,¹⁶ Y. Fukazawa,³² S. Funk,³ P. Fusco,^{14,15} F. Gargano,¹⁵ D. Gasparri,²⁰ N. Gehrels,^{21,33,34} M. Georganopoulos,²³ S. Germani,^{12,13} B. Giebels,¹⁶ N. Giglietto,^{14,15} F. Giordano,^{14,15} M. Giroletti,²³ T. Glanzman,³ G. Godfrey,³ I. A. Grenier,⁶ J. E. Grove,¹ L. Guillemot,³⁶ S. Guiriec,³⁷ Y. Hanabata,³² A. K. Harding,²¹ M. Hayashida,³ E. Hays,²¹ R. E. Hughes,¹¹ M. S. Jackson,^{28,38} G. Jóhannesson,³ A. S. Johnson,³ T. J. Johnson,^{21,34} W. N. Johnson,¹ T. Kamae,³ H. Katagiri,³² J. Kataoka,³⁹ N. Kawai,^{40,41} M. Kerr,¹⁷ J. Knödseder,⁴² M. L. Kocian,³ M. Kuss,⁵ J. Lande,³ L. Latronico,⁵ M. Lemoine-Goumard,^{30,31} F. Longo,^{7,8} F. Loparco,^{14,15} B. Lott,^{30,31} M. N. Lovellette,¹ P. Lubrano,^{12,13} G. M. Madejski,³ A. Makeev,^{1,24} M. N. Mazziotta,¹⁵ W. McConville,^{21,34} J. E. McEnery,^{21,34} C. Meurer,^{27,28} P. F. Michelson,³ W. Mitthumsiri,³ T. Mizuno,³² A. A. Moiseev,^{22,34} C. Monte,^{14,15} M. E. Monzani,³ A. Morselli,⁴³ I. V. Moskalenko,³ S. Murgia,³ P. L. Nolan,³ J. P. Norris,⁴⁴ E. Nuss,²⁵ T. Ohsugi,³² N. Omodei,⁵ E. Orlando,⁴⁵ J. F. Ormes,⁴⁴ D. Paneque,³ D. Parent,^{30,31} V. Pelassa,²⁵ M. Pepe,^{12,13} M. Pesce-Rollins,⁵ F. Piron,²⁵ T. A. Porter,⁴ S. Rainò,^{14,15} R. Rando,^{9,10} M. Razzano,⁵ S. Razzaque,^{1,2} A. Reimer,^{3,46} O. Reimer,^{3,46} T. Reposeur,^{30,31} S. Ritz,⁴ L. S. Rochester,³ A. Y. Rodríguez,¹⁸ R. W. Romani,³ M. Roth,¹⁷ F. Ryde,^{28,38} H. F.-W. Sadrozinski,⁴ R. Sambruna,²¹ D. Sanchez,¹⁶ A. Sander,¹¹ P. M. Saz Parkinson,⁴ J. D. Scargle,⁴⁷ C. Sgrò,⁵ E. J. Siskind,⁴⁸ D. A. Smith,^{30,31} P. D. Smith,¹¹ G. Spandre,⁵ P. Spinelli,^{14,15} J.-L. Starck,⁶ L. Stawarz,^{3,49} M. S. Strickman,⁵¹ D. J. Suson,⁵⁰ H. Tajima,³ H. Takahashi,³² T. Takahashi,⁵¹ T. Tanaka,³ J. B. Thayer,³ J. G. Thayer,³ D. J. Thompson,²¹ L. Tibaldo,^{6,9,10} D. F. Torres,^{18,52} G. Tosti,^{22,23} A. Tramacere,^{3,53} Y. Uchiyama,³ T. L. Usher,³ V. Vasileiou,¹⁷ N. Vilchez,⁴² V. Vitale,^{43,54} A. P. Waite,³ E. Wallace,³ P. Wang,³ B. L. Winer,¹¹ K. S. Wood,¹ T. Ylinen,^{28,38,55} M. Ziegler,⁴ M. J. Hardcastle,⁵⁶ D. Kazanas²¹

¹Space Science Division, Naval Research Laboratory, Washington, DC 20375, USA. ²National Research Council Research Associate, National Academy of Sciences, Washington, DC 20001, USA. ³W. W. Hansen Experimental Physics Laboratory, Kavli Institute for Particle Astrophysics and Cosmology, Department of Physics and SLAC National Accelerator Laboratory, Stanford University, Stanford, CA 94305, USA. ⁴Santa Cruz Institute for Particle Physics, Department of Physics and Department of Astronomy and Astrophysics, University of California at Santa Cruz, Santa Cruz, CA 95064, USA. ⁵Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, I-56127 Pisa, Italy. ⁶Laboratoire AIM, CEA-IRFU/CNRS/Université Paris Diderot, Service d'Astrophysique, CEA Saclay, 91191 Gif sur Yvette, France. ⁷Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, I-34127 Trieste, Italy. ⁸Dipartimento di Fisica, Università di Trieste, I-34127 Trieste, Italy. ⁹Istituto Nazionale di Fisica Nucleare, Sezione di Padova, I-35131 Padova, Italy. ¹⁰Dipartimento di Fisica "G. Galilei", Università di Padova, I-35131 Padova, Italy. ¹¹Department of Physics, Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH

43210, USA. ¹²Istituto Nazionale di Fisica Nucleare, Sezione di Perugia, I-06123 Perugia, Italy. ¹³Dipartimento di Fisica, Università degli Studi di Perugia, I-06123 Perugia, Italy. ¹⁴Dipartimento di Fisica "M. Merlini" dell'Università e del Politecnico di Bari, I-70126 Bari, Italy. ¹⁵Istituto Nazionale di Fisica Nucleare, Sezione di Bari, 70126 Bari, Italy. ¹⁶Laboratoire Leprince-Ringuet, École polytechnique, CNRS/IN2P3, Palaiseau, France. ¹⁷Department of Physics, University of Washington, Seattle, WA 98195–1560, USA. ¹⁸Institut de Ciències de l'Espai (IEEC-CSIC), Campus UAB, 08193 Barcelona, Spain. ¹⁹Istituto Nazionale di Astrofisica (INAF)–Istituto di Astrofisica Spaziale e Fisica Cosmica, I-20133 Milano, Italy. ²⁰Agenzia Spaziale Italiana Science Data Center, I-00044 Frascati (Roma), Italy. ²¹NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. ²²Center for Research and Exploration in Space Science and Technology and NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. ²³Department of Physics and Center for Space Sciences and Technology, University of Maryland Baltimore County, Baltimore, MD 21250, USA. ²⁴George Mason University, Fairfax, VA 22030, USA. ²⁵Laboratoire de Physique Théorique et Astroparticules, Université Montpellier 2, CNRS/IN2P3, Montpellier, France. ²⁶Department of Physics and Astronomy, Sonoma State University, Rohnert Park, CA 94928–3609, USA. ²⁷Department of Physics, Stockholm University, AlbaNova, SE-106 91 Stockholm, Sweden. ²⁸The Oskar Klein Centre for Cosmoparticle Physics, AlbaNova, SE-106 91 Stockholm, Sweden. ²⁹Dipartimento di Fisica, Università di Udine and Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, Gruppo Col-

legato di Udine, I-33100 Udine, Italy. ³⁰Université de Bordeaux, Centre d'Études Nucléaires Bordeaux Gradignan, UMR 5797, Gradignan, 33175, France. ³¹CNRS/IN2P3, Centre d'Études Nucléaires Bordeaux Gradignan, UMR 5797, Gradignan, 33175, France. ³²Department of Physical Sciences, Hiroshima University, Higashi-Hiroshima, Hiroshima 739-8526, Japan. ³³Department of Astronomy and Astrophysics, Pennsylvania State University, University Park, PA 16802, USA. ³⁴Department of Physics and Department of Astronomy, University of Maryland, College Park, MD 20742, USA. ³⁵INAF Istituto di Radioastronomia, 40129 Bologna, Italy. ³⁶Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany. ³⁷Center for Space Plasma and Aeronomic Research, University of Alabama in Huntsville, Huntsville, AL 35899, USA. ³⁸Department of Physics, Royal Institute of Technology, AlbaNova, SE-106 91 Stockholm, Sweden. ³⁹Waseda University, 1-104 Totsukamachi, Shinjuku-ku, Tokyo, 169-8050, Japan. ⁴⁰Department of Physics, Tokyo Institute of Technology, Meguro City, Tokyo 152-8551, Japan. ⁴¹Cosmic Radiation Laboratory, Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-0198, Japan. ⁴²Centre d'Étude Spatiale des Rayonnements, CNRS/UPS, BP 44346, F-30128 Toulouse Cedex 4, France. ⁴³Istituto Nazionale di Fisica Nucleare, Sezione di Roma "Tor Vergata," I-00133 Roma, Italy. ⁴⁴Department of Physics and Astronomy, University of Denver, Denver, CO 80208, USA. ⁴⁵Max-Planck Institut für Extraterrestrische Physik, 85748 Garching, Germany. ⁴⁶Institut für Astro- und Teilchenphysik und Institut für Theoretische Physik, Leopold-Franzens-Universität Innsbruck, A-6020 Innsbruck,

Austria. ⁴⁷Space Sciences Division, NASA Ames Research Center, Moffett Field, CA 94035–1000, USA. ⁴⁸NYCB Real-Time Computing, Lattingtown, NY 11560–1025, USA. ⁴⁹Astronomical Observatory, Jagiellonian University, 30-244 Kraków, Poland. ⁵⁰Department of Chemistry and Physics, Purdue University Calumet, Hammond, IN 46323–2094, USA. ⁵¹Institute of Space and Astronautical Science, Japanese Aerospace Exploration Agency, 3-1-1 Yoshinodai, Sagami-hara, Kanagawa 229-8510, Japan. ⁵²Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain. ⁵³Consorzio Interuniversitario per la Fisica Spaziale, I-10133 Torino, Italy. ⁵⁴Dipartimento di Fisica, Università di Roma "Tor Vergata," I-00133 Roma, Italy. ⁵⁵School of Pure and Applied Natural Sciences, University of Kalmar, SE-391 82 Kalmar, Sweden. ⁵⁶Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1184656/DC1
Materials and Methods

SOM Text

Figs. S1 to S3

Tables S1 to S3

References

13 November 2009; accepted 19 March 2010

Published online 1 April 2010;

10.1126/science.1184656

Include this information when citing this paper.

The Equation of State of a Low-Temperature Fermi Gas with Tunable Interactions

N. Navon,[†] S. Nascimbène,* F. Chevy, C. Salomon

Interacting fermions are ubiquitous in nature, and understanding their thermodynamics is an important problem. We measured the equation of state of a two-component ultracold Fermi gas for a wide range of interaction strengths at low temperature. A detailed comparison with theories including Monte-Carlo calculations and the Lee-Huang-Yang corrections for low-density bosonic and fermionic superfluids is presented. The low-temperature phase diagram of the spin-imbalanced gas reveals Fermi liquid behavior of the partially polarized normal phase for all but the weakest interactions. Our results provide a benchmark for many-body theories and are relevant to other fermionic systems such as the crust of neutron stars.

Recently, ultracold atomic Fermi gases have become a tool of choice to study strongly correlated quantum systems because of their high controllability, purity, and tunability of interactions (*I*). In the zero-range limit, interactions in a degenerate Fermi system with two spin-components are completely characterized by a single parameter $1/k_F a$, where a is the *s*-wave scattering length and $k_F = (6\pi^2 n)^{1/3}$ is the Fermi momentum (n is the density per spin state). In cold atom gases, the value of $|a|$ can be tuned over several orders of magnitude using a Feshbach resonance; this offers an opportunity to entirely explore the so-called BCS-BEC crossover, that is, the smooth transition from Bardeen-Cooper-Schrieffer (BCS) superfluidity at small

negative values of a to molecular Bose-Einstein Condensation (BEC) at small positive values of a (*I*, 2). Between these two well-understood limiting situations, a diverges, leading to strong quantum correlations. The description of this system is a challenge for many-body theories, as testified by the large amount of work in recent years (*I*). The physics of the BEC-BCS crossover is relevant for very different systems, ranging from neutron stars to heavy nuclei and superconductors.

In the grand-canonical ensemble and at zero temperature, dimensional analysis shows that the Equation of State (EoS) of a two-component Fermi gas, relating the pressure P to the chemical potentials μ_1 and μ_2 of the spin components can be written as

$$P(\mu_1, \mu_2, a) = P_0(\mu_1) h\left(\delta_1 \equiv \frac{\hbar}{\sqrt{2m\mu_1}a}, \eta \equiv \frac{\mu_2}{\mu_1}\right) \quad (1)$$

where $P_0(\mu_1) = 1/15\pi^2 (2m/\hbar^2)^{3/2} \mu_1^{5/2}$ is the pressure of a single-component ideal Fermi gas,

m is the atom mass, \hbar is the Planck constant divided by 2π , and δ_1 is the grand-canonical analog of the dimensionless interaction parameter $1/k_F a$. The indices 1 and 2 refer to the majority and minority spin components, respectively. From the dimensionless function $h(\delta_1, \eta)$, it is possible to deduce all the thermodynamic properties of the gas, such as the compressibility, the magnetization, or the existence of phase transitions. The aim of this paper is to measure $h(\delta_1, \eta)$ for a range of interactions (δ_1) and spin imbalances (η) and discuss its physical content. Because it contains the same information as Eq. 1, the function h will also be referred to as the EoS in the rest of the text.

In situ absorption images of harmonically trapped gases are particularly suited to investigate the EoS, as first demonstrated in (3) and (4). In the particular case of the grand-canonical ensemble, a simple formula relates the local pressure P at a distance z from the center of the trap along the z axis to the doubly integrated density profiles \bar{n}_1 and \bar{n}_2 (5).

$$P(\mu_1(z), \mu_2(z), a) = \frac{m\omega_z^2}{2\pi} (\bar{n}_1(z) + \bar{n}_2(z)) \quad (2)$$

Here, we define the local chemical potentials $\mu_i(z) = \mu_i^0 - \frac{1}{2}m\omega_z^2 z^2$, where μ_i^0 is the chemical potential of the component i at the bottom of the trap, assuming local density approximation. ω_r and ω_z are the transverse and axial angular frequencies of a cylindrically symmetric trap, respectively, and $\bar{n}_i(z) = \int n_i(x, y, z) dx dy$ is the atomic density n_i of the component i , doubly integrated over the transverse x and y directions. In a single experimental run at a given magnetic field, two images are recorded, providing $\bar{n}_1(z)$ and $\bar{n}_2(z)$ (fig. S4); the z -dependence of the chemical potentials then enables the measurement of P along a curve in the (δ_1, η) plane (6). This method was validated in (4) for the particular case of the

Laboratoire Kastler Brossel, CNRS, Université Pierre et Marie Curie, École Normale Supérieure, 24 rue Lhomond, 75231 Paris, France.

*These authors contributed equally to this work.

[†]To whom correspondence should be addressed. E-mail: navon@ens.fr



The Equation of State of a Low-Temperature Fermi Gas with Tunable Interactions

N. Navon, *et al.*

Science **328**, 729 (2010);

DOI: 10.1126/science.1187582

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/729>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/science.1187582/DC1>

This article appears in the following **subject collections**:

Physics

<http://www.sciencemag.org/cgi/collection/physics>

unitary limit $a = \infty$. Deducing the function h from the doubly integrated profiles further requires a precise calibration of ω_z and the knowledge of the central chemical potentials μ_i^0 (6).

Our experimental setup is presented in (7). We prepared an imbalanced mixture of ^6Li in the two lowest internal spin states, at the magnetic field of 834 G (where $a = \infty$), and trapped it in a hybrid magnetic-optical dipole trap. We then performed evaporative cooling by lowering the optical trap power, while the magnetic field was ramped to the final desired value for a . The cloud typically contained $N = 2$ to 10×10^4 atoms in each spin state at a temperature of $0.03(3) T_F$, justifying our $T = 0$ assumption (6). The final trap frequencies are $\omega_x/2\pi \sim 30$ Hz and $\omega_r/2\pi \sim 1$ kHz. Below a critical spin population imbalance, our atomic sample consists of a fully paired superfluid occupying the center of the trap, surrounded by a normal mixed phase and an outer rim of an ideal gas of majority component atoms (4, 7, 8).

For a given magnetic field, 10 to 20 images are taken, leading after averaging to a low-noise EoS along one line in the (δ_1, η) plane. Measurements at different magnetic fields chosen between 766 G and 981 G give a sampling of the surface $h(\delta_1, \eta)$ in the range $-1 < \delta_1 < 0.6$ and $-2 < \eta < 0.7$ (Fig. 1). Let $A(\delta_1)$ be the limiting value of the ratio of chemical potentials $\mu_1(z)/\mu_2(z)$ below which the minority density vanishes. At fixed δ_1 and $\eta < A(\delta_1)$, $h(\delta_1, \eta)$ represents the EoS of an ideal Fermi gas of majority atoms and is equal to 1. For $\eta > A(\delta_1)$, it slowly rises and corresponds to the normal mixed phase, where both spin components are present. At a critical value $\eta = \eta_c(\delta_1)$, the slope of h abruptly changes (6), the signature of a first-order phase transition from the normal phase (for $A < \eta < \eta_c$) to a superfluid phase with a lower chemical potential imbalance ($\eta > \eta_c$). We notice that the discontinuity is present for all values of δ_1 we investigated, and this feature is more pronounced on the BEC side.

Let us first consider the EoS of the superfluid phase, $\eta > \eta_c$. Each of our in situ images has, along the z axis, values of the chemical potential ratio $\eta(z) = \mu_2(z)/\mu_1(z)$ both lower and greater than η_c . In the region where $\eta(z) > \eta_c$, the doubly integrated density difference $\bar{n}_1(z) - \bar{n}_2(z)$ is constant within our signal-to-noise ratio (fig. S4). This is the signature of equal densities of the two species in the superfluid core, that is, the superfluid is fully paired. Using Gibbs-Duhem relation $n_i = \frac{\partial P}{\partial \mu_i}$, equal densities $n_1 = n_2$ imply that $P(\mu_1, \mu_2, a)$ is a function of μ and a only, where $\mu \equiv (\mu_1 + \mu_2)/2$. For the balanced superfluid, we then write the EoS symmetrically.

$$P(\mu_1, \mu_2, a) = 2P_0(\tilde{\mu})h_S\left(\tilde{\delta} \equiv \frac{h}{\sqrt{2m\tilde{\mu}a}}\right) \quad (3)$$

To avoid using negative chemical potentials, we define here $\tilde{\mu} = \mu - E_b/2$, where E_b is the molecular binding energy $E_b = -\hbar^2/ma^2$ for $a > 0$ (and 0 for $a \leq 0$). $h_S(\tilde{\delta})$ is then a single-variable function. It fully describes the ground-state

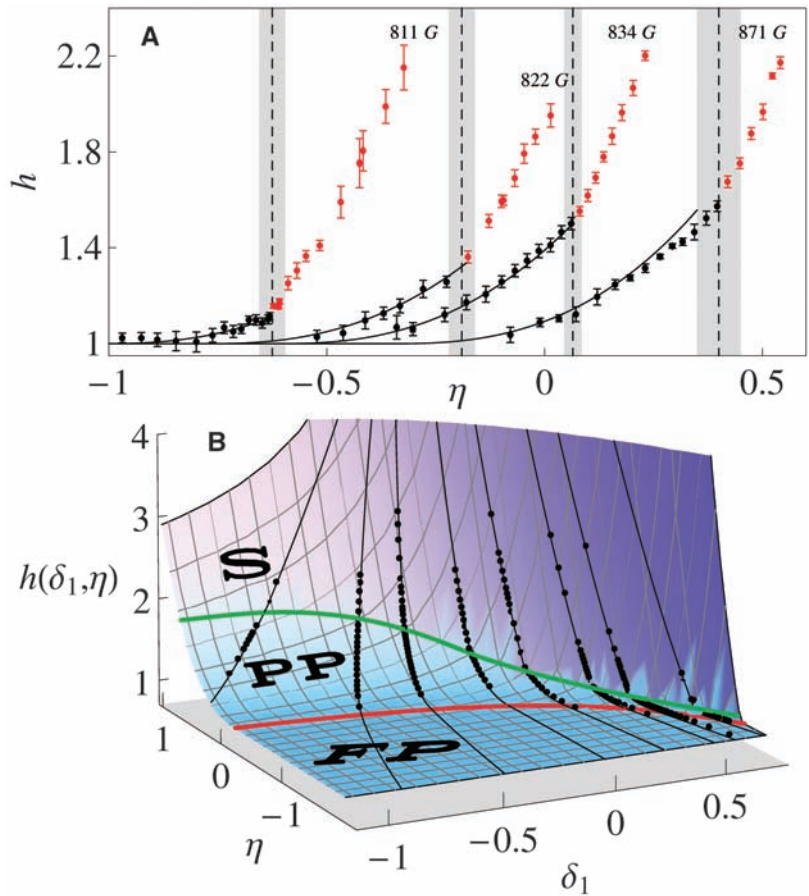


Fig. 1. $h(\delta_1, \eta)$ of a zero-temperature two-component Fermi gas in the BEC-BCS crossover. **(A)** Samples of the data for different magnetic fields. The black (red) data points correspond to the normal (superfluid) phase and are separated at $\eta_c(\delta_1)$ by a clear kink in the local slope of h . Solid black lines are the predictions of the polaron ideal gas model (Eq. 8). The scattering length corresponding to each curve is (from left to right): (1.7, 3.4, ∞ , and -1.3) in units of $10^4 a_0$, where a_0 is the Bohr radius. **(B)** $h(\delta_1, \eta)$. The black dots are data recorded for each magnetic field value (as in Fig. 1A). The black lines correspond to the parametric curves $[\delta_1(\eta), \eta]$ scanned by the density inhomogeneity in the harmonic trap (6). The red line is $A(\delta_1)$, the frontier between the fully polarized (FP) ideal gas $h = 1$ and the normal partially polarized (PP) phase. The green line is $\eta_c(\delta_1)$, marking the phase transition between the normal and superfluid (S) phases. The surface is the parametrization of $h(\delta_1, \eta)$ given in the text.

Fig. 2. $h_S(\tilde{\delta})$ of the $T = 0$ balanced superfluid in the BEC-BCS crossover (black dots). The blue solid line is the fit $h_S^{\text{BCS}}(\tilde{\delta})$ on the BCS side of the resonance; the red solid line is the fit $h_S^{\text{BEC}}(\tilde{\delta})$ on the BEC side. The dotted (dashed) red line is the mean-field (LHY) theory (32). (Inset) Zoom on the BCS side. The dotted and dashed blue lines are the EoS, including the mean-field and LHY terms, respectively. The systematic uncertainties on the x and y axes are about 5%. The errors bars represent the standard deviation of the statistical uncertainty.

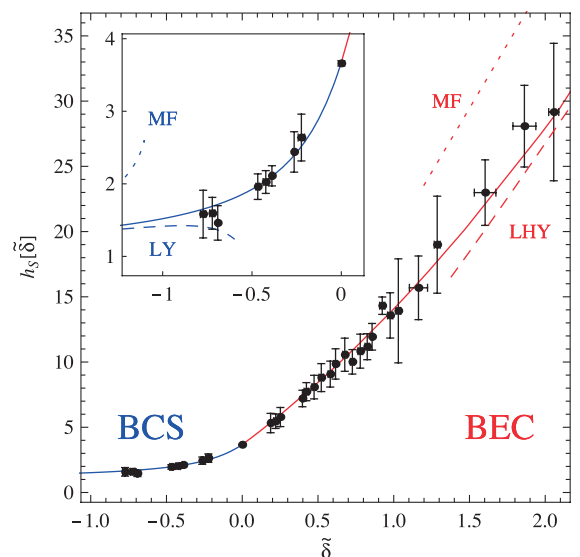
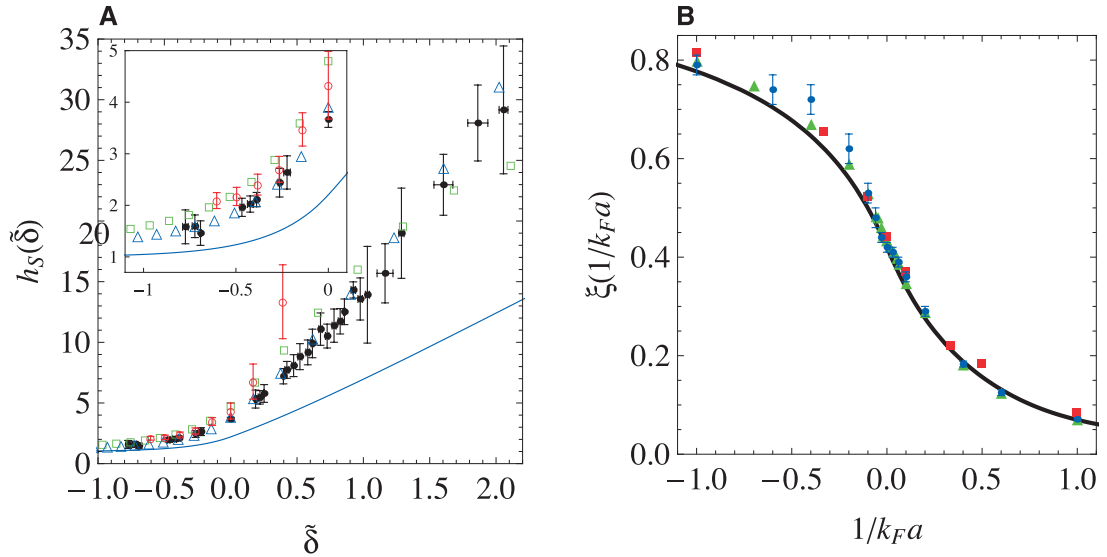


Fig. 3. Comparison with many-body theories. (A) Direct comparison of $h_S(\delta)$ with a quantum Monte-Carlo calculation [red open circles (22)], a diagrammatic method [green open squares (23)], a Nozières-Schmitt-Rink approximation [blue open triangles (21)], and the BCS mean-field theory (solid blue line). (Inset) Zoom on the BCS side. (B) EoS in the canonical ensemble $\xi(1/k_F a)$ (solid black line) deduced from the Padé-type approximants to the experimental data h_S^{BCS} and h_S^{BEC} plotted in Fig. 2. Fixed-Node Monte-Carlo theories: red squares (24), blue circles (25), and green triangles (26).



macroscopic properties of the balanced superfluid in the BEC-BCS crossover and is displayed in Fig. 2 as black dots.

To extract relevant physical quantities, such as beyond mean-field corrections, it is convenient to parametrize our data with analytic functions. In this pursuit, we use Padé-type approximants (6), interpolating between the EoS measured around unitarity and the well-known mean-field expansions on the BEC and BCS limits. The two analytic functions, h_S^{BCS} and h_S^{BEC} , are respectively represented in blue and red solid lines in Fig. 2 and represent our best estimate of the EoS in the whole BEC-BCS crossover.

On the BCS side, (δ) < 0 , h_S^{BCS} yields the following perturbative expansion of the energy in series of $k_F a$

$$E = \frac{3}{5} N E_F \left(1 + \frac{10}{9\pi} k_F a + 0.18(2)(k_F a)^2 + 0.03(2)(k_F a)^3 + \dots \right) \quad (4)$$

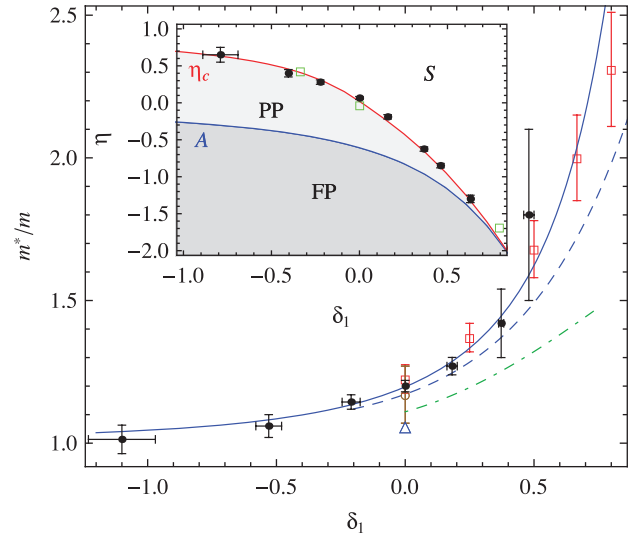
where N is the total number of atoms, E_F is the Fermi energy, and where by construction of h_S^{BCS} , the mean-field term (proportional to $k_F a$) is fixed to its exact value $10/9\pi$. We obtain beyond mean-field corrections up to the third order. The term proportional to $(k_F a)^2$ agrees with the Lee-Yang (9, 10) theoretical calculation $4(11-2\log 2)/21\pi^2 \approx 0.186$. The third-order coefficient also agrees with the value 0.030 computed in (11).

Around unitarity, the energy expansion yields

$$E = \frac{3}{5} N E_F \left(\xi_s - \zeta \frac{1}{k_F a} + \dots \right) \quad (5)$$

We find the universal parameter of the unitary $T = 0$ superfluid, $\xi_s = 0.41(1)$ with 2% accuracy. This value is in agreement with recent calculations and measurements (1). Our thermodynamic measurement $\zeta = 0.93(5)$ can be compared with a recent experimental value $\zeta = 0.91(4)$ (12), as well as the theoretical value $\zeta = 0.95$ (13), both of them

Fig. 4. Effective mass m^*/m of the polaron in the BEC-BCS crossover (black dots). The blue dashed line is a calculation from (29), red open squares (30), green dot-dashed line (26), and blue solid line (31). Measurements at unitarity through density profile analysis [blue triangle (3)] and collective modes study [brown empty circle (7)] are also displayed. (Inset) Phase diagram of a zero-temperature imbalanced Fermi gas in the BEC-BCS crossover. The blue line is the theoretical value of A (26, 29, 30) that sets the separation between the partially polarized (PP) and the fully polarized (FP) phases. Black dots are the measured values of η_c (as in Fig. 1A), which set the separation between the superfluid (S) phase and the partially polarized phase. The red line is the calculation of η_c using our EoS of the superfluid and the model (Eq. 8) for the normal phase. The green squares are lower bounds of η_c given by the values of the gap measured in (33); see (6).



obtained through the study of the pair correlation function. This experimental agreement confirms the link between the macroscopic thermodynamic properties and the microscopic short-range pair correlations, as shown theoretically in (14).

In the BEC limit, the energy of the superfluid is that of a weakly interacting Bose-Einstein condensate of molecules (9, 15)

$$E = \frac{N}{2} E_b + N \frac{\pi \hbar^2 a_{dd}}{2m} \times n \left(1 + \frac{128}{15\sqrt{\pi}} \sqrt{na_{dd}^3} + \dots \right) \quad (6)$$

where $a_{dd} = 0.6a$ is the dimer-dimer scattering length (l) and n is the dimer density. The term in $\sqrt{na_{dd}^3}$ is the well-known Lee-Huang-Yang (LHY) correction to the mean-field interaction

between molecules (9, 15). Signatures of beyond mean-field effects were previously observed through a pioneering study of collective modes (16) and density profile analysis (17), but no quantitative comparison with Eq. 6 was made. Fitting our data in the deep BEC regime with Eq. 6, we measure the bosonic LHY coefficient $4.4(5)$, in agreement with the exact value $128/15\sqrt{\pi} \approx 4.81$ calculated for elementary bosons in (9) and recently for composite bosons in (15).

Having checked this important beyond mean-field contribution, we can go one step further in the expansion. The analogy with point-like bosons suggests that the next term should be written as $[\frac{8}{3}(4\pi - 3\sqrt{3})na_{dd}^3(\log(na_{dd}^3) + B)]$ (6, 18, 19). Using $h_S^{\text{BEC}}(\delta)$ (Fig. 2) (6), we deduce the effective three-body parameter for composite bosons $B = 7(1)$. Interestingly, this value is close

to the bosonic hard-sphere calculation $B = 8.5$ (20) and to the value $B \approx 7.2$ for point-like bosons with large scattering length (19).

Our measurements also allow direct comparison with advanced many-body theories developed for homogeneous gases in the strongly correlated regime. As displayed in Fig. 3A, our data are in agreement with a Nozières-Schmitt-Rink approximation (21) but show significant differences from a quantum Monte-Carlo calculation (22) and a diagrammatic approach (23). The measured EoS strongly disfavors the prediction of BCS mean-field theory.

Comparison with Fixed-Node Monte-Carlo theories requires the calculation of the EoS $\xi(1/k_F a)$ in the canonical ensemble

$$\xi\left(\frac{1}{k_F a}\right) \equiv \frac{E - \frac{N}{2}E_b}{\frac{3}{5}NE_F} \quad (7)$$

that is deduced from $h_S^{\text{BCS}}(\tilde{\delta})$ and $h_S^{\text{BEC}}(\tilde{\delta})$ (6). As shown in Fig. 3B, the agreement with theories (24–26) is very good.

We now discuss the EoS of the partially polarized normal phase (black points in Fig. 1). At low concentrations, we expect the minority atoms to behave as noninteracting quasiparticles, the fermionic polarons (27). The polarons are dressed by the majority Fermi sea through a renormalized chemical potential $\mu_2 - A(\delta_1)\mu_1$ (28) and an effective mass $m^*(\delta_1)$ (26, 29, 30). Following a Fermi liquid picture, we propose to express the gas pressure as the sum of the Fermi pressure of the bare majority atoms and of the polarons (4).

$$h(\delta_1, \eta) = 1 + \left(\frac{m^*(\delta_1)}{m}\right)^{3/2} (\eta - A(\delta_1))^{5/2} \quad (8)$$

Our measured EoS agrees with this model at unitarity and on the BEC side of the resonance (Fig. 1), where for $m^*(\delta_1)$ we use the calculations from (30, 31). On the BCS side of the resonance, however, we observe at large minority concentrations an intriguing deviation to Eq. 8. In the BCS regime, the superfluid is less robust to spin imbalance. Consequently, the ratio of the two densities n_1/n_2 in the normal phase becomes close to unity near the superfluid/normal boundary η_c . The polaron ideal gas picture then fails.

Alternatively, we can let the effective mass m^* be a free parameter in the model in Eq. 8 in the fit of our data around $\eta = A$. We obtain the value of the polaron effective mass in the BEC-BCS crossover (Fig. 4).

An important consistency check of our study is provided by the comparison between our direct measurements of $\eta_c(\delta_1)$ (from Fig. 1, black dots in the inset of Fig. 4) and a calculated $\eta_c(\delta_1)$ from Eq. 8 and the EoS of the superfluid h_S . Assuming negligible surface tension, the normal/superfluid boundary is given by equating the pressure and chemical potential in the two phases. This procedure leads to the solid red line in the inset of Fig. 4, in excellent agreement with

the direct measurements. In addition, by integrating our measured EoS of the homogeneous gas over the trap, one retrieves the critical polarization for superfluidity of a trapped gas, in agreement with most previous measurements (6).

We have measured the equation of state of a two-component Fermi gas at zero temperature in the BEC-BCS crossover. Extensions of our work include exploring the thermodynamics of the far BEC region of the phase diagram where a new phase associated with a polarized superfluid appears (17, 26), mapping the EoS as a function of temperature, and investigating the influence of finite interaction range, which is playing a key role in higher-density parts of neutron stars.

References and Notes

- M. Inguscio, W. Ketterle, C. Salomon, Eds. *Ultra-cold Fermi Gases: Proceedings of the International School of Physics "Enrico Fermi", Course CLXIV, Varenna, 20 to 30 June 2006* (IOS Press, Amsterdam, 2008).
- A. J. Leggett, in *Modern Trends in the Theory of Condensed Matter*, A. Pekalski, R. Przystawa, Eds. (Springer-Verlag, Berlin, 1980), pp. 13–27.
- Y. Shin, *Phys. Rev. A* **77**, 041603 (2008).
- S. Nascimbène, N. Navon, K. J. Jiang, F. Chevy, C. Salomon, *Nature* **463**, 1057 (2010).
- T. Ho, Q. Zhou, *Nat. Phys.* **6**, 131 (2009).
- Materials and methods are available as supporting material on Science Online.
- S. Nascimbène et al., *Phys. Rev. Lett.* **103**, 170402 (2009).
- Y. Shin, M. Zwierlein, C. Schunck, A. Schirotzek, W. Ketterle, *Phys. Rev. Lett.* **97**, 30401 (2006).
- T. D. Lee, C. N. Yang, *Phys. Rev.* **105**, 1119 (1957).
- R. B. Diener, R. Sensarma, M. Randeria, *Phys. Rev. A* **77**, 023626 (2008).
- G. Baker Jr., *Rev. Mod. Phys.* **43**, 479 (1971).
- H. Hu et al., "Universal structure of a strongly interacting Fermi superfluid," <http://arxiv.org/abs/1001.3200> (2010).
- C. Lobo, I. Carusotto, S. Giorgini, A. Recati, S. Stringari, *Phys. Rev. Lett.* **97**, 100405 (2006).
- S. Tan, *Ann. Phys.* **323**, 2971 (2008).
- X. Leyronas, R. Combescot, *Phys. Rev. Lett.* **99**, 170402 (2007).
- A. Altmeyer et al., *Phys. Rev. Lett.* **98**, 040401 (2007).
- Y. I. Shin, A. Schirotzek, C. H. Schunck, W. Ketterle, *Phys. Rev. Lett.* **101**, 070404 (2008).
- T. Wu, *Phys. Rev.* **115**, 1390 (1959).
- E. Braaten, H. W. Hammer, T. Mehen, *Phys. Rev. Lett.* **88**, 040401 (2002).
- S. Tan, *Phys. Rev. A* **78**, 013636 (2008).
- H. Hu, X. Liu, P. Drummond, *Europhys. Lett.* **74**, 574 (2006).
- A. Bulgac, J. Drut, P. Magierski, *Phys. Rev. A* **78**, 023625 (2008).
- R. Haussmann, W. Rantner, S. Cerrito, W. Zwerger, *Phys. Rev. A* **75**, 023610 (2007).
- S. Chang, V. Pandharipande, J. Carlson, K. Schmidt, *Phys. Rev. A* **70**, 043602 (2004).
- G. E. Astrakharchik, J. Boronat, J. Casulleras, A. S. Giorgini, *Phys. Rev. Lett.* **93**, 200404 (2004).
- S. Pilati, S. Giorgini, *Phys. Rev. Lett.* **100**, 030401 (2008).
- C. Lobo, A. Recati, S. Giorgini, S. Stringari, *Phys. Rev. Lett.* **97**, 200403 (2006).
- A. Schirotzek, C.-H. Wu, A. Sommer, M. W. Zwierlein, *Phys. Rev. Lett.* **102**, 230402 (2009).
- R. Combescot, A. Recati, C. Lobo, F. Chevy, *Phys. Rev. Lett.* **98**, 180402 (2007).
- N. Prokof'ev, B. Svistunov, *Phys. Rev. B* **77**, 020408 (2008).
- R. Combescot, S. Giraud, X. Leyronas, *Europhys. Lett.* **88**, 60007 (2009).
- In the BEC limit, the grand-canonical EoS expands as $h_S(\tilde{\delta}) \approx 15\pi a/4a_{dd}\tilde{\delta} - 16\sqrt{2}$; the first term is the mean-field interaction and the second is the LHY correction asymptotic behavior.
- A. Schirotzek, Y. I. Shin, C. H. Schunck, W. Ketterle, *Phys. Rev. Lett.* **101**, 140403 (2008).
- We thank K. Jiang for participation in the early phase of the experimental work. We are grateful to X. Leyronas, C. Mora, Y. Castin, F. Werner, R. Combescot, J. Dalibard, F. Gerbier, and G. Shlyapnikov for stimulating discussions and critical comments on the manuscript. We thank S. Giorgini, P. Drummond, J. Drut, R. Haussmann, and W. Zwerger for providing us with their data. We acknowledge support from European Research Council, European Science Foundation (Euroquam), SCALA (Scalable Quantum Computing with Light and Atoms), Agence Nationale de la Recherche FABIOLA (Fermions and Bosons in Optical Lattices), Région Ile de France Institut Francilien de Recherche sur les Atomes Froids, and Institut Universitaire de France.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1187582/DC1
Materials and Methods

Figs. S1 to S4

References

27 January 2010; accepted 30 March 2010

Published online 15 April 2010;

10.1126/science.1187582

Include this information when citing this paper.

Nanoscale Three-Dimensional Patterning of Molecular Resists by Scanning Probes

David Pires,¹ James L. Hedrick,² Anuja De Silva,³ Jane Frommer,² Bernd Gotsmann,¹ Heiko Wolf,¹ Michel Despont,¹ Urs Duerig,¹ Armin W. Knoll^{1*}

For patterning organic resists, optical and electron beam lithography are the most established methods; however, at resolutions below 30 nanometers, inherent problems result from unwanted exposure of the resist in nearby areas. We present a scanning probe lithography method based on the local desorption of a glassy organic resist by a heatable probe. We demonstrate patterning at a half pitch down to 15 nanometers without proximity corrections and with throughputs approaching those of Gaussian electron beam lithography at similar resolution. These patterns can be transferred to other substrates, and material can be removed in successive steps in order to fabricate complex three-dimensional structures.

To date, a wide variety of techniques has been available for nanofabrication (1), including electron beam lithography (EBL)

and scanning probe lithography (SPL) (2–4) as direct-write methods. Although EBL is used in critical applications such as the fabrication of

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/732>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/science.1187851/DC1>

This article appears in the following **subject collections**:

Materials Science

http://www.sciencemag.org/cgi/collection/mat_sci

masks for optical lithography (5), the fabrication of patterns at resolutions below 30 nm with EBL is difficult because of proximity effects (6). The throughput also scales unfavorably as the resolution is increased (7–9). Scanned probes are inherently capable of addressing surfaces with subnanometer precision. This capability has been exploited to create patterns with atomic resolution with a scanning tunneling microscope (STM) (10), albeit at extremely low patterning speeds.

In SPL, the close proximity of the tip to the substrate opens up multiple ways to modify the sample by the tip at resolutions below 50 nm (2–4). We demonstrate the use of a heated tip to locally desorb material from a thin film of an organic molecular glass. Structuring of organic materials as resists is our preferred strategy, because it follows the same philosophy as wafer-scale lithography methods in complementary metal-oxide semiconductor technology and enables the transfer of the pattern into arbitrary underlying substrates (e.g., by etching techniques). The approach avoids large stresses on the tip, such as are needed for plowing (11) or ultrasonic removal of material (12), and also mitigates the problem of piling up of excess material inherent in these approaches. Moreover, it does not require a conductive substrate in close proximity to the tip, as is necessary for current-induced depolymerization (13), field-induced deposition (14, 15), or field emission from the tip (16, 17).

Heated tips have already been used for patterning resists by removing covalently bound material (18, 19) and by inducing local chemical modifications (20, 21). However, cleavage of a chemical bond by a heated tip at very fast time scales is difficult because of the large energetic barriers of covalent bonds. Instead, we chose a resist material in which organic molecules are bound by hydrogen-bonding interactions into a glassy bulk state. The H bonds provide sufficient stability to the material for imaging and processing, but are sufficiently weak to be efficiently thermally activated by the hot tip.

The molecular structure of the resist used here is shown in Fig. 1C (22). Similar molecules have been designed as high-resolution photoresists (23, 22), and their relatively low molecular weight enables physical vapor deposition (PVD) of the material (24). The material under study has a molecular weight of 715 g/mol and a PVD temperature of $\sim 220^\circ\text{C}$. Six hydroxy groups are located at the periphery of the molecule and give rise to numerous hydrogen-bonding interactions in the bulk of the material, as inferred from the high glass transition temperature T_g of 126°C .

The patterning principle (25) is shown in Fig. 1, A and B. The structure to be written was defined

by a pixel set. At each pixel of the programmed pattern (red outline in Fig. 1, A and B), a force and a temperature stimulus were applied to the cantilever, which pulled the tip into contact from its rest position of ~ 300 nm above the surface (Fig. 1A) and induced the evaporation of a controlled amount of organic material (Fig. 1B). The heat and the force stimuli were provided by voltage pulses V_h and V_f , applied to a resistive heater coupled to the tip and to the substrate for electrostatic actuation, respectively. For a given interaction time (i.e., pulse duration), the depth of a written pixel depended on the applied force and temperature, as shown in Fig. 1D for pulse durations of 5.5 μs . The temperature given refers to the heater temperature T_{heater} . The actual increase relative to room temperature (RT) in resist temperature below the tip apex was $(T_{\text{resist}} - RT) \approx 0.4$ to $0.5 (T_{\text{heater}} - RT)$ (26).

Two temperature regimes could be identified that were separated by a marked change in the sensitivity of pixel formation on the applied force. Below a threshold of $\sim 330^\circ\text{C}$, the force-temperature characteristics were as expected for hot embossing (26, 27). For a given indent depth, the graphs are linear and converge to a common

intersection point at $\sim 350^\circ\text{C}$, as expected for a material with a T_g of 126°C (18). The indentation characteristics changed drastically if a heater temperature greater than $\sim 330^\circ\text{C}$ was applied. This transition marks the softening temperature of the material at a microsecond time scale and the crossover from an embossing regime to a regime of effective material removal (Fig. 1D). Above this threshold, sufficient thermal energy is provided to the material for breaking the hydrogen bonds at the microsecond time scale, allowing the molecular constituents to become mobile and escape from the surface. Experimental evidence supports a model in which the molecules diffuse along the tip into hotter tip regions, where the molecules are evaporated into the environment (25).

Figure 2 summarizes the results achieved using the patterning process. For the pattern shown in Fig. 2A, a total of 3.6×10^4 pixels were written at a pitch of 29 nm, resulting in uniformly recessed structures of 8 ± 1 nm depth, as evidenced by the depth histogram shown in Fig. 2C. The quality and uniformity of the patterning process can be seen in Fig. 2B and fig. S4. Each pixel of the programmed pattern, indicated by the red line, is well resolved. The depth of the pixels

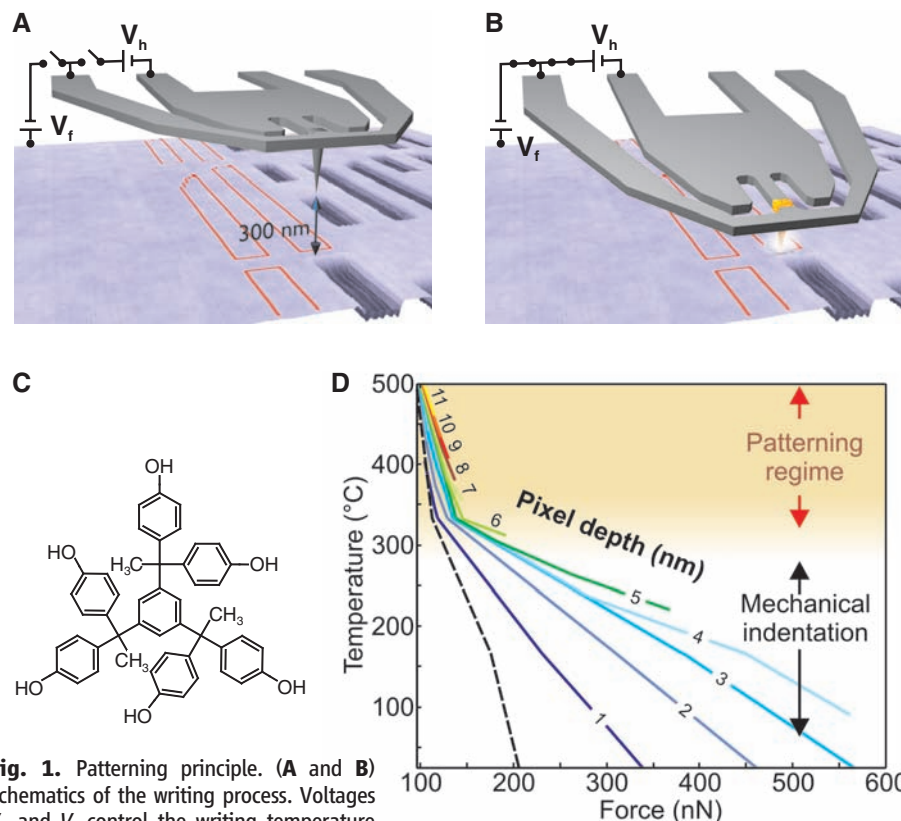
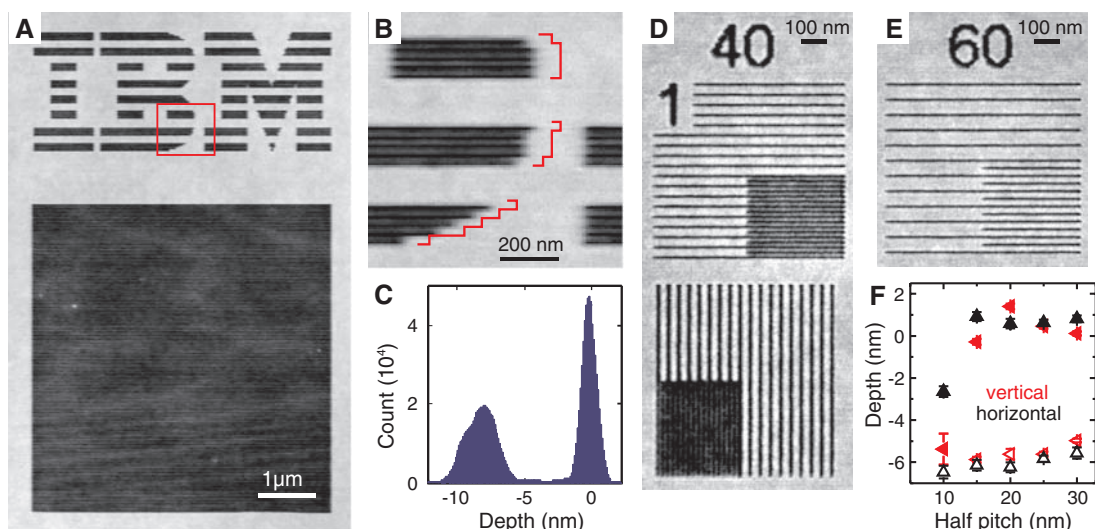


Fig. 1. Patterning principle. (A and B) Schematics of the writing process. Voltages V_h and V_f control the writing temperature and the electrostatic force, respectively. If no voltage is applied (A), the tip rests 300 nm above the surface. A pixel in the programmed bitmap (red outline) is written by simultaneously applying a force and temperature pulse for several microseconds. The force pulse pulls the tip into contact while the heat pulse heats the tip and triggers the patterning process. (C) Molecular structure of the phenolic compound used as resist. (D) Equi-depth lines of the pixels created in the resist upon single exposure events as a function of applied temperature and force. The dashed line indicates the writing threshold, determined by extrapolation to zero depth. Two regimes are separated by a threshold temperature of $\sim 330^\circ\text{C}$. Below the threshold, the resist is mechanically deformed; above the threshold, it is efficiently removed.

¹IBM Research–Zurich, Säumerstrasse 4, 8803 Rüschlikon, Switzerland. ²IBM Research–Almaden, 650 Harry Road, San Jose, CA 95120, USA. ³IBM Research–Watson, T. J. Watson Research Center, Yorktown Heights, NY 10598, USA.

*To whom correspondence should be addressed. E-mail: ark@zurich.ibm.com

Fig. 2. Patterning of the molecular glass resist. (A) Topographic image of a pattern written into the molecular glass resist using a pixel size of 29 nm. Force and heat-pulse durations of $5.5\ \mu\text{s}$, a tip heater temperature of $300^\circ \pm 30^\circ\text{C}$, and a force of $80 \pm 10\ \text{nN}$ were applied for each pixel. (B) Close-up of the red box shown in (A), demonstrating the feature quality of the writing process. The red outline corresponds to the shape of the programmed image and is shifted relative to the real pattern for clarity. (C) Depth histogram of (A). The patterning depth is $8 \pm 1\ \text{nm}$. (D) A field of dense lines at a line pitch of 40 nm written into a molecular glass film $20 \pm 2\ \text{nm}$ thick. One quadrant is written at half the line pitch (i.e., at 20 nm pitch). A pixel pitch of 10 nm, a tip heater temperature of $500^\circ \pm 30^\circ\text{C}$, and a force of $100 \pm 10\ \text{nN}$ were used. (E) Same as the top panel



is the same regardless of the existence of written or unwritten neighboring pixels. In other words, at a pitch of 29 nm, proximity effects were absent. In the patterning process, a total volume of $\sim 0.2\ \mu\text{m}^3$ was removed from the surface, which is more than 10^5 times that of the tip apex of $\sim 1000\ \text{nm}^3$. However, no traces of displaced or redeposited material were detected either on the sample or at the tip. The efficiency of the process suggests that the molecules are removed without breaking covalent bonds and creating reactive species, which typically lead to tip contamination (25).

The achievable resolution increased with decreasing patterning depth. Figure 2, D to F, shows the result of high-resolution patterning at a depth of $\sim 6\ \text{nm}$. A dense line pattern written at varying pitch was chosen to study the transition from single-line patterning to complete removal of material within a given area. For this experiment, pre- and post-heating of the tip for 2 and 3 μs , respectively, were applied at a force pulse duration of $5\ \mu\text{s}$ to reduce thermal expansion and contraction of the cantilever structure during a patterning event. The average depth of the single lines and the average height of the topography between adjacent lines are shown in Fig. 2F. The error bars correspond to standard variations of the respective values. The depth of the lines has a constant value of $\sim 6\ \text{nm}$, independent of the distance to neighboring lines. Down to a half pitch of 15 nm, the height of the topography between the lines was unaffected at the level of the unwritten resist surface; that is, any interference or proximity effects of neighboring lines were absent. Decreasing the half pitch further to 10 nm led to a reduction of the interline topography and enabled patterning of a flat, recessed area having a depth corresponding to the depth of the single lines (Fig. 2D). If we define the resolution of the process as the half pitch of fully separated dense

of (D) at a pitch of 60 nm (one quadrant is written at 30 nm pitch). (F) Depth of the written lines (open symbols) and height of the topography as measured between adjacent lines (solid symbols).

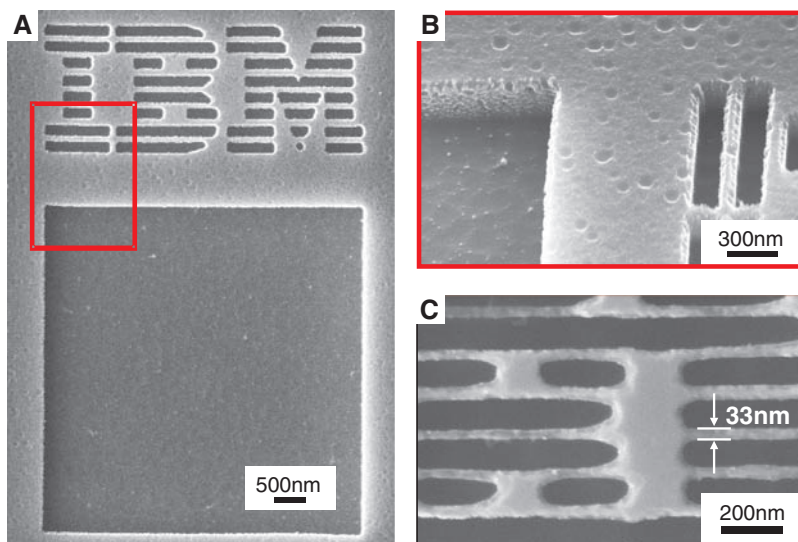


Fig. 3. Pattern transfer into silicon. (A) SEM image of the pattern shown in Fig. 2 transferred 400 nm deep into silicon. (B) Tilted view of the structure indicated by the red box in (A). (C) Zoom into a similar structure as shown in (A) but written at half the pixel pitch. The smallest lines fabricated in silicon had a width of $\sim 30\ \text{nm}$.

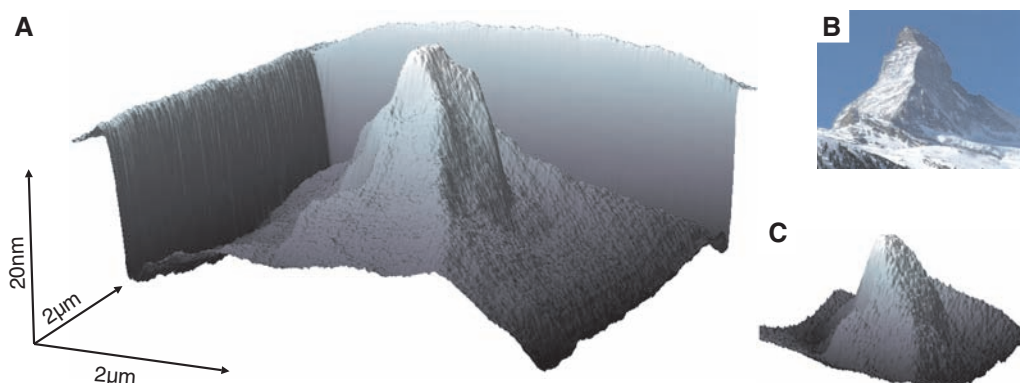
lines, a resolution of $\sim 15\ \text{nm}$ was achieved using a tip with an apex radius of $\sim 5\ \text{nm}$. The resolution can be further increased by using sharper tips or at the cost of a shallower pattern formation. In general, a finite value of the patterning depth is required to enable further processing—for example, for transferring the pattern into the underlying substrate or hard mask as described below.

For device fabrication, the structure depth in the target substrate must substantially exceed the depth of the patterns described above. For example, according to the International Technology Roadmap for Semiconductors, for the 22-nm node, a pattern in a resist 45 to 80 nm thick is required (28). A hard-mask strategy (25) has been adopted to amplify the pattern first into a

resist layer of adequate thickness and subsequently into the silicon substrate. In short, the 8-nm-deep pattern within the molecular glass is first transferred into a 3-nm-thick silicon oxide hard mask by means of reactive ion etching (RIE). Subsequently, the hard mask serves as an etch-selective layer for the pattern transfer 70 nm deep into a poly(styrene-*r*-benzocyclobutene) (PS-BCB) (29) transfer layer.

At this stage, the required amplification into a thicker polymer layer was achieved, and different strategies could be implemented for further processing, such as deposition and liftoff of metal layers. Here we chose to transfer the pattern into the silicon substrate by another RIE process (25). Shown in Fig. 3, A and B, are

Fig. 4. Three-dimensional patterning. **(A)** AFM scan of the replica of the Matterhorn written into the molecular glass (3D data source: geodata © swisstopo). The structure was written using 120 steps of layer-by-layer removal. **(B)** Photograph of the Matterhorn in Switzerland (photographer: Marcel Wiesweg; source: Wikimedia). **(C)** AFM image of a Matterhorn replica transferred into silicon.



scanning electron microscopy (SEM) images of the final pattern in silicon as obtained after transfer of the pattern shown in Fig. 2A. The depth of the final structure in silicon is enhanced by a factor of 50 and amounts to more than 400 nm.

During the etch process, the pattern quality was maintained and only a fine rounding of the edges in the IBM logo was observed, caused by a slight over-etching. The large amplification enables the fabrication of high-aspect-ratio structures from the low-aspect-ratio SPL master. The same pattern as shown in Fig. 2A was written at half the pixel pitch (i.e., 15 nm), and a part of the final pattern in silicon is shown in Fig. 3C. The thickness of the walls between the patterned areas is 33 nm in the final silicon structure. As a result, vertical walls 33 nm thick and ~400 nm tall were fabricated, corresponding to an aspect ratio of >10 (30).

Material removal can be done in stages so that three-dimensional structures can be fabricated, as shown in Fig. 4. A given three-dimensional structure is processed into the resist by removing successive layers of defined thickness from the same sample area. A pyramid was created (fig. S6) with linear side walls, sharp edges, and a pointed tip. It demonstrates that each patterning step is reproducible and independent of the already existing structures created in preceding steps. By exploiting these properties, highly detailed and complex shapes can be reproduced. We fabricated a nanoscale replica of the Swiss mountain Matterhorn (Fig. 4B) in a 100-nm-thick molecular glass film (Fig. 4A). The patterning was achieved in 120 steps, resulting in a structure 25 nm tall (see also movie S1). Fine details of the original are reproduced in the nanometer-scale replica. The conformal reproduction of the original proves that the final structure is a linear superposition of well-defined single patterning steps. The result of a direct transfer of a similar structure into the underlying substrate is shown in Fig. 4C. Using a mixture of SF_6 and C_4F_8 as etch gas provides a height amplification of the structures into silicon by a factor of ~3.

Areal throughputs in the range of $5 \times 10^4 \mu\text{m}^2/\text{hour}$ are achievable today assuming a duty cycle of 50%, which is sufficient for rapid proto-

typing applications in nanotechnology and approaches the throughput of EBL at similar resolution (7). High parallelization of a similar process using the same levers has been developed for a thermomechanical probe data storage system (31), enabling a potential throughput enhancement proportional to the number of levers operated at the same time. Assuming 1000 levers in parallel leads to throughput numbers that cannot be achieved by any other direct-write technology to date at the same resolution. At the same time, the direct development method offers in situ inspection and, potentially, in situ repair of structures not fully reproduced. Proximity effects are absent at pixel dimensions of ~15 nm and higher. Therefore, no computationally elaborate, substrate-dependent adjustment of the exposure dose is needed. Moreover, the dry nature of the entire patterning process eliminates solvent-related problems such as swelling-induced instability or drying-induced collapse of dense structures. Furthermore, the ability to generate three-dimensional templates having a nanometer-precise profile is a complementary match to nanoimprint lithography (32).

References and Notes

- Y. Xia, J. A. Rogers, K. E. Paul, G. M. Whitesides, *Chem. Rev.* **99**, 1823 (1999).
- R. Garcia, R. V. Martinez, J. Martinez, *Chem. Soc. Rev.* **35**, 29 (2006).
- A. A. Tseng, A. Notargiacomo, T. P. Chen, *J. Vac. Sci. Technol. B* **23**, 877 (2005).
- D. Wouters, U. S. Schubert, *Angew. Chem. Int. Ed.* **43**, 2480 (2004).
- H. J. Levinson, *Principles of Lithography* (SPIE—International Society for Optical Engineering, Bellingham, WA, 2005).
- T. H. P. Chang, *J. Vac. Sci. Technol.* **12**, 1271 (1975).
- C. R. K. Marrian, D. M. Tennant, *J. Vac. Sci. Technol. A* **21**, S207 (2003).
- R. Menon, A. Patel, D. Gil, H. I. Smith, *Mater. Today* **8**, 26 (2005).
- J. A. Liddle, G. M. Gallatin, L. E. Ocola, in *Three-Dimensional Nanoengineered Assemblies* (Materials Research Society Symposium Proceedings, vol. 739), T. M. Orlando, L. Merhari, K. Ikuta, D. P. Taylor, Eds. (Materials Research Society, Warrendale, PA, 2003), pp. 19–30.
- D. M. Egler, E. K. Schweizer, *Nature* **344**, 524 (1990).
- L. L. Sohn, R. L. Willett, *Appl. Phys. Lett.* **67**, 1552 (1995).
- F. Iwata, T. Matsumoto, R. Ogawa, A. Sasaki, *J. Vac. Sci. Technol. B* **17**, 2452 (1999).
- C. Martín, G. Rius, X. Borrísé, F. Pérez-Murano, *Nanotechnology* **16**, 1016 (2005).
- R. V. Martínez, N. S. Losilla, J. Martinez, Y. Huttel, R. Garcia, *Nano Lett.* **7**, 1846 (2007).
- M. Rolandi, I. Suez, A. Scholl, J. M. Frechet, *Angew. Chem. Int. Ed.* **46**, 7477 (2007).
- S. W. Park, H. T. Soh, C. F. Quate, S. I. Park, *Appl. Phys. Lett.* **67**, 2415 (1995).
- K. Wilder, C. F. Quate, B. Singh, D. F. Kyser, *J. Vac. Sci. Technol. B* **16**, 3864 (1998).
- B. Gotsmann, U. Duerig, J. Frommer, C. J. Hawker, *Adv. Funct. Mater.* **16**, 1499 (2006).
- A. A. Milner, K. Zhang, Y. Prior, *Nano Lett.* **8**, 2017 (2008).
- O. Fenwick et al., *Nat. Nanotechnol.* **4**, 664 (2009).
- R. Szożkiewicz et al., *Nano Lett.* **7**, 1064 (2007).
- A. De Silva et al., *Chem. Mater.* **20**, 1606 (2008).
- J. Dai, S. W. Chang, A. Hamad, N. Felix, C. K. Ober, *Chem. Mater.* **18**, 3404 (2006).
- F. Pfeiffer, N. M. Felix, C. Neuber, C. K. Ober, H. W. Schmidt, *Adv. Funct. Mater.* **17**, 2336 (2007).
- See supporting material on Science Online.
- B. Gotsmann, M. A. Lantz, A. Knoll, U. Duerig, *Nanotechnology*, Vol. 6: *Nanoprobes* (VCH, Weinheim, Germany, 2009), pp. 121–169.
- The resist temperature at a heater temperature of 350°C is ~170°C. The temperature difference from the glass transition temperature of $T_g = 126^\circ\text{C}$ is consistent with the time-temperature behavior of glassy organic materials. Preliminary experiments recording the indent depth as a function of heater temperature and writing time reveal a decrease of the writing threshold temperature with increasing exposure duration by ~20°C per decade at microsecond time scales, leveling off at a heater temperature of ~250°C for writing times of seconds.
- International Technology Roadmap for Semiconductors, 2008 Update (www.itrs.net/Links/2008ITRS/Update/2008Tables_FOCUS_B.xls).
- E. Harth et al., *J. Am. Chem. Soc.* **124**, 8653 (2002).
- Because of the high aspect ratio, we could not measure the depth by atomic force microscopy (AFM). Therefore, we assume that the etch depth for these narrow trenches is similar to that for the large structures of 400 nm.
- A. Pantazi et al., *IBM J. Res. Develop.* **52**, 493 (2008).
- K. Watanabe et al., *J. Vac. Sci. Technol. B* **22**, 22 (2004).
- We gratefully acknowledge the invaluable support from the probe storage team at the IBM Research Laboratory in Rüschlikon, in particular U. Drechsler for the pattern transfer and the fabrication of thermomechanical probe sensors, and P. Seidler, W. Riess, R. Miller, and R. Allen for stimulating discussions.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1187851/DC1
Materials and Methods
SOM Text
Figs. S1 to S6
Movie S1

2 February 2010; accepted 30 March 2010
Published online 22 April 2010;
10.1126/science.1187851
Include this information when citing this paper.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
<http://www.sciencemag.org/cgi/content/full/328/5979/736>

Supporting Online Material can be found at:
<http://www.sciencemag.org/cgi/content/full/328/5979/736/DC1>

This article **cites 39 articles**, 1 of which can be accessed for free:
<http://www.sciencemag.org/cgi/content/full/328/5979/736#otherarticles>

This article appears in the following **subject collections**:
Physics, Applied
http://www.sciencemag.org/cgi/collection/app_physics

Visualizing the Electron Scattering Force in Nanostructures

Chenggang Tao,* W. G. Cullen, E. D. Williams†

In nanoscale metal wires, electrical current can cause structural changes through electromigration, in which the momentum of electrons biases atomic motion, but the microscopic details are complex. Using in situ scanning tunneling microscopy, we examined the effects of thermally excited defects on the current-biased displacement of monatomic islands of radius 2 to 50 nanometers on single-crystal Ag(111). The islands move opposite to the current direction, with velocity varying inversely with radius. The force is thus in the same direction as electron flow and acts on atomic defect sites at the island edge. The unexpectedly large force on the boundary atoms can be decreased by over a factor of 10 by adding a mildly electron-withdrawing adsorbate, C_{60} , which also modifies the step geometry. The low coordination of the identified scattering sites is the likely origin of the large force.

When electrical current passes through a wire, the current carriers can scatter off of interfaces or defects and transfer momentum to atoms in these regions, resulting in biased diffusion of mobile atoms. This electromigration force is so small that it may result in a directional bias in fewer than one in 10 million diffusion steps, but the cumulative effects generate micrometer-scale voids in the small metal wires in modern electronics (1, 2). At the nanoscale, electromigration forces may be equally destructive but also provide opportunities for control of structure as in nanogap formation (3), for driving dopants into nanowires (4), and for coupling electrical signals to atomic fluctuations (5, 6). We used scanning tunneling microscopy (STM) to observe the biased displacement of metallic islands in the presence of an applied current (Fig. 1A) so as to correlate the magnitude of the electromigration force to specific atomic structures, in this case “kink” sites at step edges. Using a step-continuum analysis, we quantified the site-resolved atomic force and found that the electromigration effects are much stronger than had been expected theoretically for non-kinked step edges.

Observations of mass diffusion due to electromigration under extremely high current densities ($>10^{13}$ A/m²) have been presented for nano-gap formation (7, 8); however, the complex geometries involved are not amenable to extracting quantitative information by modeling mass flow. The alternative approach of using theoretical understanding to quantify electron scattering at surface and interface defect sites requires an evolution from classical descriptions of bulk electromigration. The classical approaches use mean-field methods and must be reframed in terms of atomic structures, their local electron density, and nonequilibrium transport characteristics in order to understand electron scattering and electromigration in nanoscale structures.

In this study, the direct observation of the biased displacement of nanoscale structures produced a quantitative value for the electromigration force that can be identified with a specific type of structural defect, independent of assumptions about the bulk electrical properties. The test systems consist of silver (Ag) islands containing between 100 and 100,000 atoms. Elec-

trical current (defined, as is standard, as being in the opposite direction to electron flow) through the substrate provides momentum transfer to the atoms in the islands, yielding a response that depends on the diffusion mechanism (9). Because surface diffusion is easiest for free atoms on terraces (D_T) or via thermally excited atomic-scale kinks at island edges (D_L), island displacement is most likely to occur via atomic attachment to or motion along the island edges. As a result, island motion does involve a shifting center of mass (COM) but not the net simultaneous displacement of all the atoms in the island.

The electromigration force can move an island via atomic motion in two different ways (Fig. 1B). Freely diffusing atoms on the terrace experience a drift velocity ($v_{EM} = D_T F / k_B T$, where F is the electromigration force, k_B is the Boltzmann constant, and T is temperature), creating a net flux of atoms onto one side of the island and a corresponding flux away from the opposite side. The resulting net displacement of the island's COM will be opposite the electromigration force, and (for a given electromigration force) the velocity of the island will be determined by the atomic flux per unit length of the island edge and

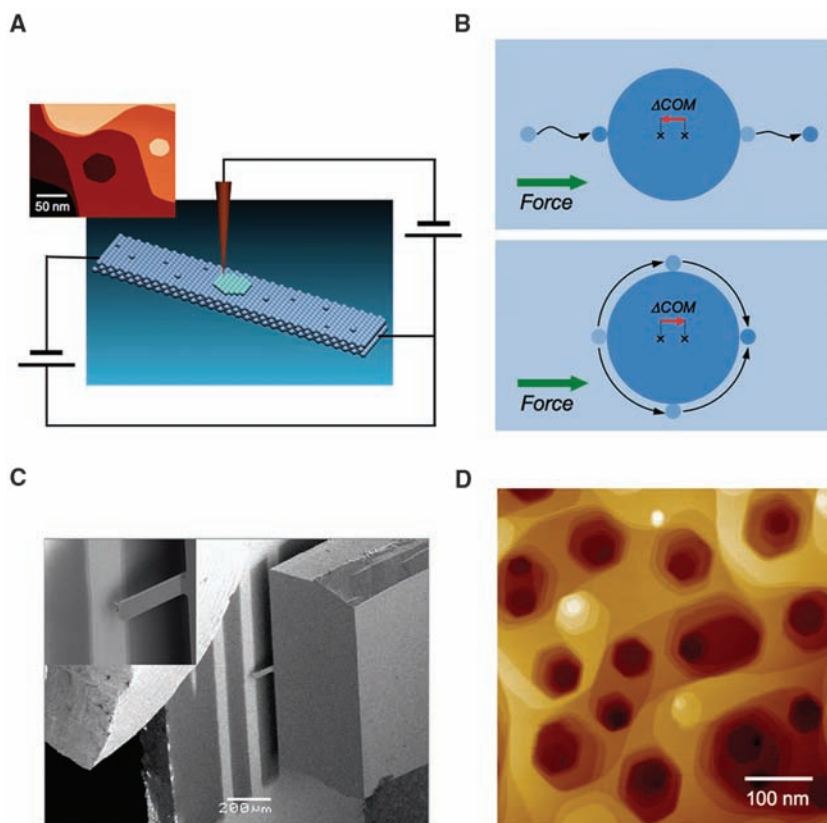


Fig. 1. (A) Schematic drawing of the experimental setup. (B) Schematic illustration of two mechanisms (top, attachment/detachment; bottom, edge diffusion); whichever mechanism biases the displacement of mass from one end of an island to the other shifts the COM. The starting positions of atoms are light blue; the end positions are dark blue. (C) SEM image of the experimental setup: Ag stripes and scanning tip. In the middle of the image, bright stripes are Ag stripes. STM images are obtained over the central part of the middle stripe as shown in the inset, which is a zoom-in SEM image of the STM tip and Ag stripe. (D) STM topography image of the surface Ag stripes after sputtering procedure. The image size is (500 × 500) nm², and the tunneling current is 52 pA with a sample bias of −1.87 V.

Materials Research Science and Engineering Center and Department of Physics, University of Maryland, College Park, MD 20742–4111, USA.

*Present address: Department of Physics, University of California at Berkeley, Berkeley, CA 94720, USA.

†Present address: BP, 1 St. James Square, London SW1Y 4PD, UK.

is thus independent of island size. Alternatively, the island will move if the force preferentially moves atoms along its edges from the “upwind” side to the “downwind” side. In this case, the displacement of the COM is opposite to that of terrace diffusion case. For edge diffusion, the island displacement is parallel to the electromigration force, and (for a given electromigration force) the COM velocity is determined by the island radius. A combination of the direction of island displacement and quantitative dependence of the rate of displacement on the island size yields both the mechanism and the magnitude of the force (9).

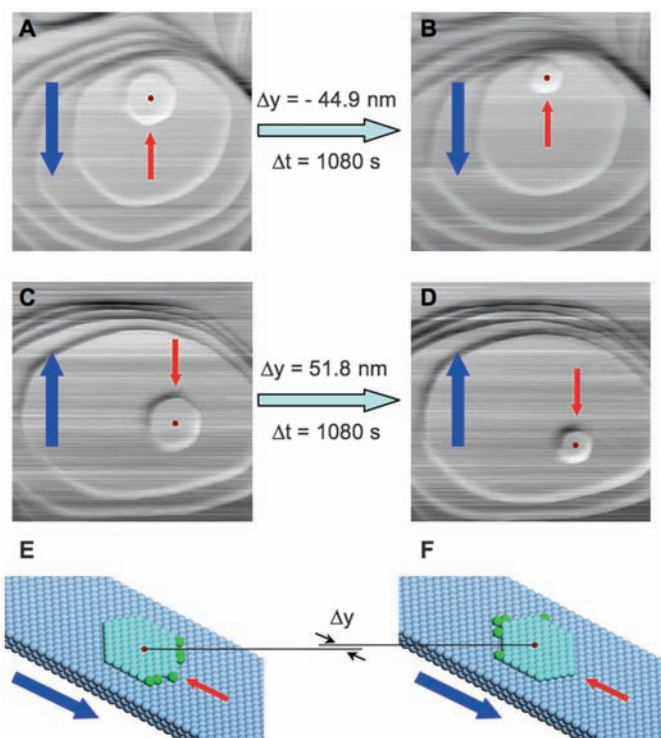
To study these two cases, it was necessary to maximize the current density j while minimizing Joule heating, which otherwise drives the rates of nanostructure evolution past the point of observability by STM. The Ag thin film is shadow mask-patterned in stripes, with a width of ~ 50 μm and thickness of ~ 350 nm (Fig. 1C). The stacks of adatom islands or vacancy islands range in height (or depth) from five to eight monatomic layers (Fig. 1D). Starting with this surface morphology, we applied direct current to the stripes. After thermal stabilization (about 30 to 40 min), we located the STM tip onto the middle area of the stripe and then measured the displacement of monatomic islands with the direct current still applied. Additional precautions relevant for these STM measurements are discussed in (10). For the step-bending measurements, samples were also prepared with

precisely controlled C_{60} decoration of the step edges (11).

At a current density of 6.7×10^9 A/m² and a sample temperature of 318 K (10), the biased displacement can be observed by comparing sequential STM images, as in the still frames, from movie S1, shown in Fig. 2.

These results are quantified in Fig. 3, in which the COM of the islands is shown as a function of time for current applied in the positive and negative y directions. The initial position of the front edge of each island is shown by the horizontal line. The biased displacement is demonstrated by the strong anisotropy of the displacements in the parallel and the perpendicular directions (y and x), which is dramatically different from Brownian motion of islands without a biasing current (12). In the perpendicular (x) direction, there is no net displacement. In the parallel (y) direction, the measured displacement is up to 90 nm, limited by the time for the islands to totally decay (~ 1700 s). The island displacement is parallel to the direction of electron flow. During the measurements, the island area decreases linearly with time (Fig. 3B), which is similar to the linear decay behavior of isolated islands (13) and different from the nonlinear decay of islands bounded by three-dimensional stacks of concentric terraces (14, 15). The island decay provides a natural method for evaluating the dependence of the drift velocity on island size, for comparison with theoretical prediction (9).

Fig. 2. Biased displacement of monolayer islands driven by electromigration force at $T = 318$ K. Shown for comparison are image pairs for monolayer islands that were measured in exactly the same area but with different current directions. In (A) and (B), in which the current direction is downward, the island displacement is upward, whereas the island displacement is downward in (C) and (D), in which the current direction is upward. The shrinking island size seen in both pairs of images is a thermal effect; the same shrinking occurs at this temperature without the electrical current present. The blue arrows indicate the direct current direction, and the orange arrows indicate the direction of electron flow. The image size is (500×500) nm². (E) and (F) schematically illustrate the island displacement that is due to the force on edge atoms, in which the blue atoms are the substrate, the lighter blue atoms are those in the bulk of the island, and the green atoms are diffusing edge atoms of the island. The small dark red dot indicates the COM of the monolayer island in (A) to (F).



From the displacement measurements, the drift velocity of the islands was calculated by averaging the slopes of two adjacent position-time points without smoothing. The range of island velocities is from 0.02 nm/s to 0.12 nm/s, which can be compared with the step velocity at copper (Cu) grain boundaries of ~ 5 nm/s at approximately four-times-larger current density (16). The correlation of the velocities with the island radii is presented in Fig. 3C and shows a strong velocity-radius dependence. The black solid line is a one-parameter fit of the experimental data to $v = C/R$, where C is a physical constant (Eq. 1) and R is the island radius, yielding $C = 1.93 \pm 0.09$ nm²/s. An inverse radius dependence of the velocity has been predicted for island displacement biased by motion of atoms around the edges of the island (9) in contrast with the constant velocity predicted for a pure attachment/detachment mechanism. An edge-diffusion mechanism is also consistent with previous measurements (17) of island edge and step fluctuations on Ag(111), which have shown that the dominant mode of Ag mass transport is via edge diffusion in this temperature range. Lastly, the measured decay rate [$da/dt = -B$, where B is a physical constant (10)] and COM velocity ($v = C/R$) for the circular islands was combined and integrated to yield the time-dependent positions of the front and back edges of the island, as shown by the solid curves in Fig. 3D. These results confirm the mechanism of island displacement as step-edge diffusion and thus the origin of the electromigration force as electron collisions.

The expression for the island drift velocity resulting from atomic diffusion around the edges is (9)

$$v = -\frac{aD_L F}{k_B T R} \quad (1)$$

where a is the lattice constant. Given the measured product of v and R above, and assuming an isotropic island edge, F can be calculated given the edge diffusion constant D_L . Previous measurements of step fluctuation dynamics (18) and calculations of the island-edge free energy (19) show that the value of D_L depends strongly on the orientation of the island edge (10). The edge diffusion constant for high-symmetry (straight) step edges is $D_{L,s} = 1.8 \times 10^4$ nm²/s, whereas diffusion on low-symmetry (rough) step edges is much more difficult, with an average value of $D_{L,r} = 4.1 \times 10^2$ nm²/s, with uncertainty (1 SD) of $\sim 40\%$ on the values. The actual island shape consists of straight edges joined by rounded corners. A lower limit for F can be obtained by assuming the largest diffusion constant, that of the straight edges, yielding $|F_{\min}| = 0.012$ meV/nm ($|F_{\min}|/j_{\text{bulk}} = 1.8 \times 10^{-4}$ eV cm/A, where j_{bulk} is the bulk current density). However, the diffusing atoms must traverse the rough regions of the island edge in series with the straight regions, suggesting that a higher value of F is needed to generate the measured island speed. Analysis of

the island shape (10) shows that ~27% of the island edge is rough. Calculating an effective diffusion constant as $1/D_{L,eff} = 0.73/D_{L,s} + 0.27/D_{L,r}$ gives $D_{L,eff} = 1.4 \times 10^3 \text{ nm}^2/\text{s}$, and the corresponding F magnitude is $0.15 \pm 0.06 \text{ meV/nm}$. Normalizing this to the bulk current density gives $|F/j_{bulk}| = 2.2 \times 10^{-3} \text{ eV cm/A}$.

This value is strikingly larger than the expected force on a step on Ag(111), which should be approximately one and a half to two times as large as that for a Ag adatom on Ag(111) (20), for which $|F/j_{bulk}| = 3.0 \times 10^{-5} \text{ eV cm/A}$. It is also larger than calculated values in other systems, such as for vacancy migration in Cu (21), for which $|F/j_{bulk}| = 6.6 \times 10^{-5} \text{ eV cm/A}$, or the scattering force in a single-atom-wide chain of silicon (Si) atoms, for which $|F/j_{bulk}| \cong 1 \times 10^{-4} \text{ eV cm/A}$ (22). Instead, the magnitude of F/j_{bulk} is comparable with calculations for ions in carbon nanotubes (4) gated into strong conductance and is consistent with estimates that are based on experimental measurement of the electromigration-biased fluctuations of steps on Ag(111) (23). These values are shown in Table 1.

We modified the electronic environment of the step edge (24) by means of adsorption of C_{60} , which is known to withdraw electrons from Ag (25). C_{60} decorates step edges, and qualitative observation of the morphology of the decorated surface shows dramatic changes under the same current density ($6.7 \times 10^9 \text{ A/m}^2$) (10). The decorated steps change orientation, with the steps moving from the perpendicular to parallel to the current direction during observation. To quantify this effect, we identified pinned steps in which step curvature is reproducibly altered by the electromigration force (Fig. 4A). The arrow in each image indicates the current direction. The step bends against the current direction, and thus in the direction of the electron flow, so that the decorated step looks like a sail bowing under the wind. Similarly to the equilibrium (no current)-decorated island edges (11), the decorated steps here also fluctuate. To obtain the steady-state configuration (Fig. 4B), more than 10 images were averaged to quantify each configuration.

The changes in step shape are caused by displacement of Ag atoms under the electromigration force, which is equivalent to an imposed gradient in chemical potential (26)

$$F = -d\mu/dx \tag{2}$$

where F is applied in the x direction and μ is the local chemical potential. Atomic motion under this chemical potential gradient will be opposed by the Gibbs-Thomson effect, which is the change in step chemical potential attributable to curvature. On (111) surfaces, the explicit relation between the step chemical potential and the curvature κ is

$$\mu = -\frac{\sqrt{3}}{2} \tilde{\beta} a^2 \kappa \tag{3}$$

where $\tilde{\beta}$ is the step stiffness. From our previous study on decorated island fluctuations, we have

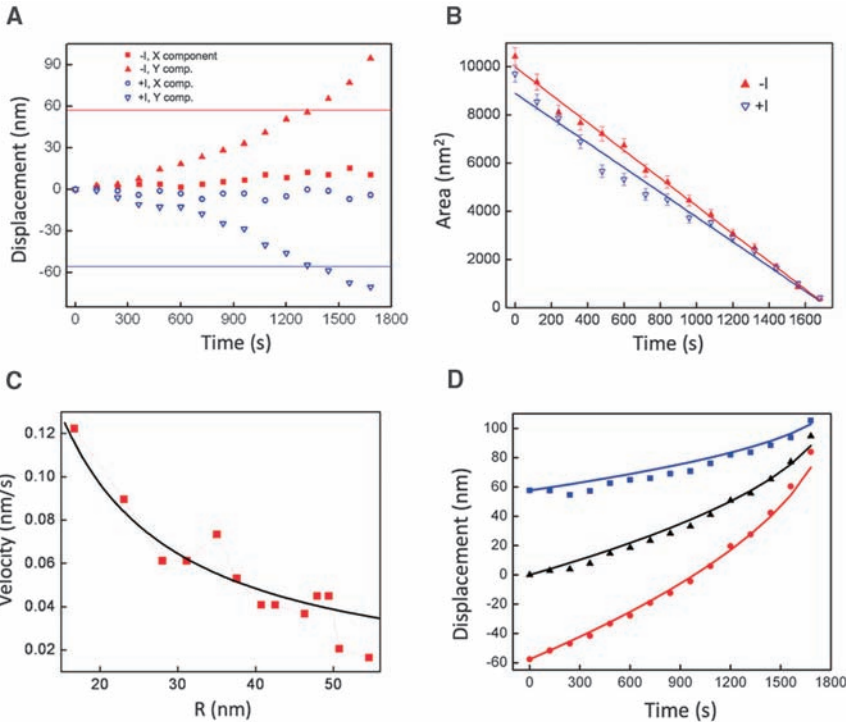


Fig. 3. (A) The displacement of island mass centers as a function of time. Here the y axis is parallel to the current direction, and the x axis is perpendicular to the current direction. The red and blue data ($-I$ and $+I$) correspond to current in the y - and y + directions, respectively. The horizontal lines indicate the initial position of the front edge of each island. (B) Island areas as a function of time. The solid lines are linear fits and yield the decay rate $-5.76 \pm 0.11 \text{ (nm}^2/\text{s)}$ for the $-I$ direction island and $-5.13 \pm 0.20 \text{ (nm}^2/\text{s)}$ for the $+I$ direction island. (C) Drift velocity as a function of the island radius. The solid black line is one parameter fit, $v = C/R$. The best fitting parameter is $C = 1.93 \pm 0.09 \text{ (nm}^2/\text{s)}$. (D) The position of the front edge (squares), COM (triangles), and back edge (circles) of the $-I$ direction island. The curves are calculated by integrating the two relationships $dA/dt = -B$ and $v = C/R$ (Eq. 1) and using the measured values of $B = 5.8 \text{ nm}^2/\text{s}$ and $C = 1.9 \text{ nm}^2/\text{s}$.

Table 1. Comparison of representative values of the electromigration force, shown using several commonly used presentations: F/j_{bulk} , $z^* = Fe/j_{bulk}\rho_{bulk}$, and $\sigma_{eff} = Felmv_F j_{bulk}$, where e is the magnitude of the electron charge. The values shown for the clean island edge include the effects of edge roughness. The lower limits obtained by neglecting edge roughness are 1/12.5 times smaller.

System	$ F/j_{bulk} $ (eV cm/A)	$ z^* $	$\sigma_{eff} \text{ (nm}^2\text{)}$	Reference
Vacancy in bulk Cu, calculated	6.6×10^{-5}	39	1.2×10^{-2}	(20)
Ion in gated carbon nanotube, calculated	$\sim 1.3 \times 10^{-3}$			(4)
Atomic Si wire, calculated	$\sim 1.1 \times 10^{-4}$			(21)
Ag atom on Ag(111), calculated	3.0×10^{-5}	17	6.1×10^{-3}	(19)
Step edge on Al, calculated		43		(27)
Step edge on Ag(111), measured		85 to 255 (lower limit)	0.03 to 0.09	(22)
Island edge site on Ag(111), measured	$(2.2 \pm 0.9) \times 10^{-3}$	1260 ± 500	4.5×10^{-1}	This work
C_{60} -decorated step edge on Ag(111), measured	1.9×10^{-4}	111	3.9×10^{-2}	This work

$\tilde{\beta} = 0.65 \text{ eV/nm}$ (11, 27). Given the smooth step shape, we can determine the local curvature of the step configurations. The solid lines in Fig. 4B are polynomial fits to the experimental data, from

which we can calculate the curvature along the step. The resulting force magnitude is $F = 1.30 \times 10^{-2} \text{ meV/nm}$ ($|F/j_{bulk}| = 1.9 \times 10^{-4} \text{ eV cm/A}$), which is obtained by averaging the two config-

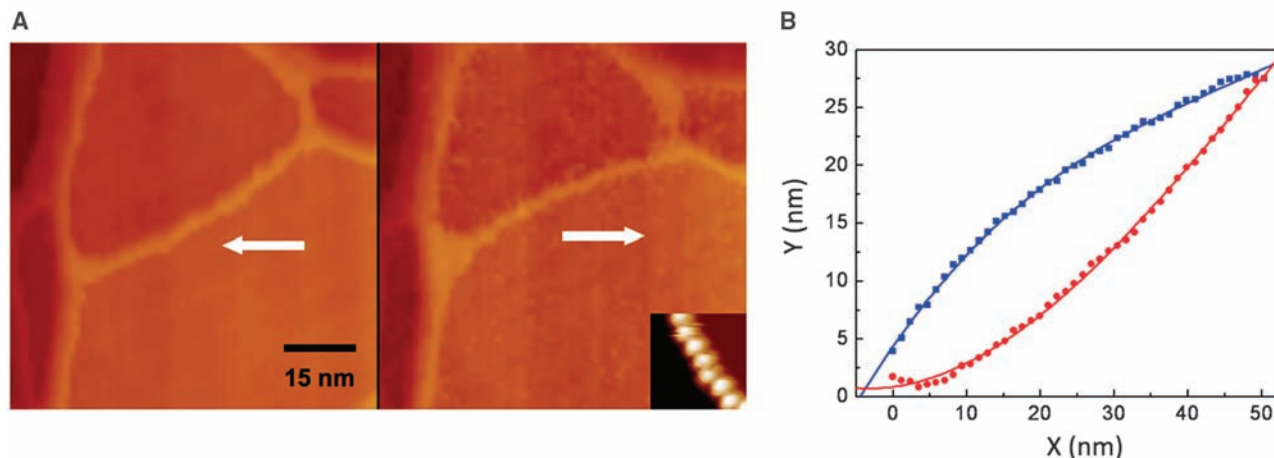


Fig. 4. (A) STM images of the steady-state configuration of a C_{60} -decorated step under applied current in two opposing directions. (Inset) Molecular structure of the chain of C_{60} molecules. (B) Quantified configuration of the

decorated steps in (A). Each configuration is obtained by averaging more than 10 images. The red circles and the blue squares are for the left and right configurations shown in (A), respectively.

urations. So, the force acting on the decorated steps is comparable with the lower limit and smaller, by slightly more than 1 SD, than the best value determined for the electromigration wind force on the bare island edges. However, it is still substantially larger than calculated values for electromigration in metals, and in particular about a factor of 4 larger than the expected value for a clean step on Ag(111).

Unlike macroscopic measurements of electromigration effects, the forces measured here are identified with a specific atomic structure—in this case, that of kink sites (for example, positions where thermal excitation has displaced an atom from the low-energy configuration) at a step. This provides an opportunity to understand this force from first principles. In the classical description, the surprisingly large magnitude in this case can arise from at least three effects: current crowding, changes in scattering cross section, and changes in local electron density. Current crowding, the variation of local current density near constrictions in flow path, can cause local increases of up to about a factor of 2 in extreme geometries (21); however, the constriction due to current flow around the island edge is likely to be more modest. The steps of the edge of the island themselves can act as a reflecting barrier, increasing the cross section for scattering as compared with a free atom on the surface by about a factor of at least 1.5 for Al(001) (28). Lastly, atomic relaxation at step edges causes redistribution of local charge density (29, 30), with especially strong enhancement—up to a factor of 10 on Cu(532)—near kink sites (24). Surface states near steps could influence the magnitude of the electromigration force by increasing or decreasing the electron density near the Fermi level. However, a reported unoccupied state on Ag(111) occurs well above the Fermi level (31) and thus should not affect the behavior of the electrons involved in transport, which are near the Fermi level.

The combination of these three effects therefore may account for up to a factor of 20 increase

in the force on kink-site atoms at the island edge as compared with that of a free Ag atom. Using the value for Ag/Ag(111) in Table 1, this yields a force magnitude up to $\sim 10^{-3}$ eV cm/A, which is slightly more than 1 SD below the measured value of $(2.2 \pm 0.9) \times 10^{-3}$ eV cm/A for the bare island edges. The measured force decreases with the addition of C_{60} , possibly through a reduction in local valence electron density because of charge transfer to the C_{60} or as a result of changes in the local kink geometry. Because the kink scattering site is much smaller than the mean free path of the transport electrons, the effects described above should be understood more directly in the context of Landauer ballistic conduction (32), coupled with nonequilibrium transport calculations (33). In Landauer ballistic conduction, the local current density rather than the bulk electric field defines the scattering environment, and there is a local potential drop at each scatterer. The corresponding dipole was first observed directly by using scanning tunneling potentiometric (STP) imaging of 5- to 10-nm-scale defects on a thin bismuth (Bi) film under current bias (34), and a more recent observation of a Ag film on Si has associated potential drops with specific step resistivities (35). First-principles calculations of the electronic structure at defect sites within a nonequilibrium transport formalism should yield the transmission coefficients for the defect sites directly, allowing a full understanding of the variability of the electromigration force with atomic scale structure.

The mechanisms proposed here to explain the large electromigration force are generally applicable to defects at surfaces and interfaces. Thus, we predict that the large forces revealed in this atomically specific measurement of the effects of electron scattering at island-edge kink sites will be a common nanoelectronic effect that is due to ubiquitous interfacial defect structures at the nanoscale (3, 16, 36–38). As a result, interesting structural transformations because of kinetic instabilities, previously thought to require onerously large current densities (39), may be

more generally accessible. The enhanced electron scattering will also couple into nanoelectronic transport characteristics, with implications for reliability, noise, switching, and possibly stochastic resonance. These and more exotic possibilities, such as harnessing the electromigration force for work (40), will be predictable as continuing experimental and theoretical work leads to a full understanding of the relation between atomic structure and scattering strength.

References and Notes

- P. S. Ho, T. Kwok, *Rep. Prog. Phys.* **52**, 301 (1989).
- H. Yasunaga, A. Natori, *Surf. Sci. Rep.* **15**, 205 (1992).
- H. Park, A. K. L. Lim, A. P. Alivisatos, J. Park, P. L. McEuen, *Appl. Phys. Lett.* **75**, 301 (1999).
- S. Heinze, N.-P. Wang, J. Tersoff, *Phys. Rev. Lett.* **95**, 186802 (2005).
- O. Pierre-Louis, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **76**, 062601 (2007).
- E. D. Williams *et al.*, *N. J. Phys.* **9**, 387 (2007).
- T. Taychatanapat, K. I. Bolotin, F. Kuemmeth, D. C. Ralph, *Nano Lett.* **7**, 652 (2007).
- T. Kizuka, H. Aoki, *Appl. Phys. Express* **2**, 075003 (2009).
- O. Pierre-Louis, T. L. Einstein, *Phys. Rev. B* **62**, 13697 (2000).
- Materials and methods are available as supporting material on Science Online.
- C. Tao, T. J. Stasevich, W. G. Cullen, T. L. Einstein, E. D. Williams, *Nano Lett.* **7**, 1495 (2007).
- K. Morgenstern, E. Laegsgaard, F. Besenbacher, *Phys. Rev. Lett.* **86**, 5739 (2001).
- K. Morgenstern *et al.*, *Phys. Rev. B* **63**, 045412 (2001).
- K. Thurmer *et al.*, *Phys. Rev. Lett.* **87**, 186102 (2001).
- A. Ichimiya *et al.*, *Phys. Rev. Lett.* **84**, 3662 (2000).
- K. C. Chen, W. W. Wu, C. N. Liao, L. J. Chen, K. N. Tu, *Science* **321**, 1066 (2008).
- M. Giesen, *Prog. Surf. Sci.* **68**, 1 (2001).
- C. Tao, T. J. Stasevich, T. L. Einstein, E. D. Williams, *Phys. Rev. B* **73**, 125436 (2006).
- T. J. Stasevich *et al.*, *Phys. Rev. B* **71**, 245414 (2005).
- M. F. G. Hedouin, P. J. Rous, *Phys. Rev. B* **62**, 8473 (2000).
- J. Hoekstra, A. P. Sutton, T. N. Todorov, A. Horsfield, *Phys. Rev. B* **62**, 8568 (2000).
- Z. Q. Yang, M. Di Ventra, *Phys. Rev. B* **67**, 161311 (2003).
- O. Bondarchuk *et al.*, *Phys. Rev. Lett.* **99**, 206801 (2007).
- F. Mehmood, A. Kara, T. S. Rahman, *Surf. Sci.* **600**, 4501 (2006).
- L. Wang, H. Cheng, *Phys. Rev. B* **69**, 165417 (2004).
- K. Thurmer, D. J. Liu, E. D. Williams, J. D. Weeks, *Phys. Rev. Lett.* **83**, 5531 (1999).

27. T. J. Stasevich, C. G. Tao, W. G. Cullen, E. D. Williams, T. L. Einstein, *Phys. Rev. Lett.* **102**, 085501 (2009).
28. P. J. Rous, *Phys. Rev. B* **59**, 7719 (1999).
29. Y. N. Mo, W. G. Zhu, E. Kaxiras, Z. Y. Zhang, *Phys. Rev. Lett.* **101**, 216101 (2008).
30. J. Y. Park *et al.*, *Phys. Rev. Lett.* **95**, 136802 (2005).
31. A. Spert *et al.*, *Phys. Rev. B* **77**, 085422 (2008).
32. R. S. Sorbello, *Solid State Phys.* **51**, 159 (1997).
33. J. Taylor, H. Guo, J. Wang, *Phys. Rev. B* **63**, 245407 (2001).
34. B. G. Briner, R. M. Feenstra, T. P. Chin, J. M. Woodall, *Phys. Rev. B* **54**, R5283 (1996).
35. J. Homoth *et al.*, *Nano Lett.* **9**, 1588 (2009).
36. Y. Zhang, W. van Dongen, B. He, T. Block, C. Tegenkamp, *Appl. Phys. Lett.* **89**, 223903 (2006).
37. C. Jin, K. Suenaga, S. Iijima, *Nano Lett.* **8**, 1127 (2008).
38. T. Sun *et al.*, *Phys. Rev. B* **79**, 041402 (2009).
39. M. Rusanen, P. Kuhn, J. Krug, *Phys. Rev. B* **74**, 245423 (2006).
40. D. Dundas, E. J. McEniry, T. N. Todorov, *Nat. Nanotechnol.* **4**, 99 (2009).
41. This work has been supported by the University of Maryland NSF Materials Research Science and Engineering Center under grant DMR 05-20471, including use of the Shared

Experimental Facilities. Infrastructure support is also provided by the University of Maryland NanoCenter and the Center for Nanophysics and Advanced Materials.

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/736/DC1

Materials and Methods

Figs. S1 to S4

References

Movie S1

4 January 2010; accepted 7 April 2010

10.1126/science.1186648

Viscosity of MgSiO₃ Liquid at Earth's Mantle Conditions: Implications for an Early Magma Ocean

Bijaya B. Karki^{1*} and Lars P. Stixrude²

Understanding the chemical and thermal evolution of Earth requires knowledge of transport properties of silicate melts at high pressure and high temperature. Here, first-principles molecular dynamics simulations show that the viscosity of MgSiO₃ liquid varies by two orders of magnitude over the mantle pressure regime. Addition of water systematically lowers the viscosity, consistent with enhanced structural depolymerization. The combined effects of pressure and temperature along model geotherms lead to a 10-fold increase in viscosity with depth from the surface to the base of the mantle. Based on these calculations, efficient heat flux from a deep magma ocean may have exceeded the incoming solar flux early in Earth's history.

Silicate liquids likely played a crucial role in terrestrial mass and heat transport in Earth's history. Molten silicates would have controlled the dynamics of the predicted magma ocean [a largely or completely molten mantle that is expected during Earth's earliest stages (1)] and continue to influence the transport of modern magmas at the present. If such a magma ocean existed, the rates of initial thermal evolution (via convection) and chemical evolution (via crystal settling and melt percolation) of Earth's interior would be primarily controlled by the melt viscosity (2). The ability of melts to carry xenoliths from great depths in the mantle (3) also depends on the melt viscosity, in addition to melt composition. Moreover, melts are considered to be responsible for the ultralow velocity zone (ULVZ) in the deep mantle detected by seismology (4, 5).

Despite their importance, transport properties, including the viscosity of molten silicates, are unknown over almost the entire mantle pressure regime, which reaches 136 GPa at the core-mantle boundary. Because of experimental difficulties, the viscosity of MgSiO₃ liquid, the dominant composition of Earth's mantle, has only been measured at ambient pressures (6). In fact, viscosity measurements of any silicate melts have been limited to relatively low pressures

(<13 GPa) (7–12). In many silicate liquids, the viscosity depends non-monotonically on pressure over the range that has been measured, making extrapolations highly uncertain. Theoretical computations serve as a complementary approach. Previous calculations were primarily based on atomistic models (13–15), which permit much faster computation but have the disadvantage of being based on empirical force fields, the forms of which are uncertain. On the other hand, the first-principles approach is more robust because it makes no assumptions about the nature of bonding or the shape of the charge density and

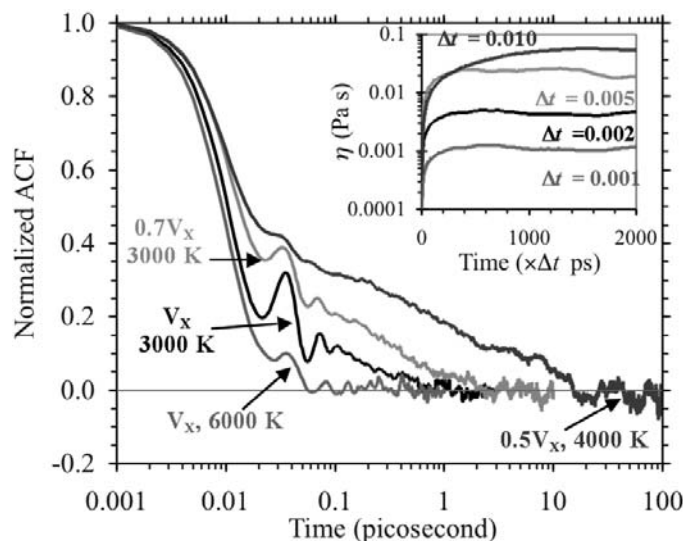
is thus in principle equally applicable to the study of a wide variety of materials problems, including liquids. We previously calculated the structure and thermodynamic properties of MgSiO₃ and MgSiO₃-H₂O liquids from first principles (16, 17), finding good agreement with extant experimental data over the entire mantle pressure-temperature regime. Unlike these equilibrium properties, the transport properties such as viscosity require much longer simulation (18).

Here, we determine the viscosity of two key liquids over the entire mantle pressure regime from density functional theory (18). MgSiO₃ serves as an analog composition for a magma ocean, whereas MgSiO₃-H₂O liquid allows us to explore the role of melt composition, focusing on H₂O as the component that is known to have the largest influence on the viscosity at low pressure (19). The shear viscosity (η) was calculated by using the Green-Kubo relation

$$\eta = \frac{V}{3k_B T} \int_0^\infty \left\langle \sum_{i < j} \sigma_{ij}(t + t_0) \cdot \sigma_{ij}(t_0) \right\rangle dt \quad (1)$$

where σ_{ij} (i and $j = x, y, z$) is the stress tensor, which is computed directly at every time step of the simulation, V is volume, k_B is the Boltzmann constant, T is temperature, t is time, and t_0 represents the time origin. The shear-stress autocorrelation function (the integrand of Eq. 1) decays to zero more slowly at lower temperature and higher pressure, requiring longer simulation

Fig. 1. Time convergences of the calculated stress autocorrelation function (ACF) and viscosity (inset) of MgSiO₃ melt (without water) at different conditions. The run durations are 18 ps (V_x , 6000 K, 7.5 GPa), 60 ps (V_x , 3000 K, 1.8 GPa), 72 ps ($0.7V_x$, 3000 K, 25 GPa), and 172 ps ($0.5V_x$, 4000 K, 135 GPa), where V_x is the reference volume ($38.9 \text{ cm}^3 \text{ mol}^{-1}$).



¹Department of Computer Science, Department of Geology and Geophysics, Louisiana State University, Baton Rouge, LA 70803, USA. ²Department of Earth Sciences, University College London, London WC1E 6BT, UK.

*To whom correspondence should be addressed. E-mail: karki@csc.lsu.edu

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/740>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/328/5979/740/DC1>

This article **cites 31 articles**, 9 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/328/5979/740#otherarticles>

This article appears in the following **subject collections**:

Geochemistry, Geophysics

http://www.sciencemag.org/cgi/collection/geochem_phys

runs (Fig. 1). We find that the integral values, and hence the computed viscosity, converge over time intervals much shorter than the total simulation durations (Fig. 1, inset). The fact that the shear stress autocorrelation function decays to zero within the time scale of our simulations means that the Maxwell relaxation time of silicate liquids (20) remains much shorter than seismic periods over the entire mantle regime and that seismic wave propagation through melts that may exist in the ULVZ will occur in the relaxed limit. We further confirm that the simulated system is in the liquid state at each pressure-temperature condition by examining the mean-square displacements (fig. S1) and radial distribution functions (fig. S2). Our approach is expected to be more robust than the commonly

used indirect approach of estimating the melt viscosity from the self-diffusion coefficient via the classic Eyring relation (20, 21). The validity of the Eyring relation as applied to silicate liquids has been questioned on the basis of experiments (11).

Over most of the pressure range of our investigation, viscosity increases with increasing pressure (Fig. 2A). The calculated viscosity increases by a factor of ~ 140 for anhydrous silicate melt over the entire mantle pressure regime at 4000 K. The activation volume $V_{\eta}^* = (d \ln \eta / dP)_T$ varies systematically over most of this range, tending to decrease with increasing pressure and with increasing temperature. At the lowest temperature and pressure, we find that the viscosity behaves anomalously:

decreasing with increasing pressure initially, reaching a minimum value near 5 GPa at 3000 K, and then increasing on further compression. Low-pressure experimental studies have found viscosity decreasing with increasing pressure in highly polymerized silicate melts (8, 11, 23). We attribute the initial decrease in viscosity with increasing pressure to the presence of fivefold coordinated silicon, which acts as a transition state accommodating viscous flow (16, 24). The variation of viscosity in this anomalous regime is small compared with the total variation in viscosity over the mantle pressure-temperature range. The calculated viscosities show large and systematic deviations from Arrhenian behavior (Fig. 2B). The activation energy decreases with increasing temperature, consistent with the behavior of moderately fragile liquids (25).

Silicate melt with 10 weight percent H_2O is two to four times less viscous than the anhydrous melt at all pressure-temperature conditions studied (Fig. 2). We have previously shown that the self-diffusion coefficients of the hydrous liquid are systematically higher than those of the anhydrous liquid (26). The region of anomalous pressure dependence of the viscosity and diffusion are weak or absent in the case of hydrous silicate liquid. Our first-principles results confirm that the dynamical enhancement (smaller viscosity and larger diffusivity) occurs in hydrous silicate liquid because water systematically depolymerizes the melt structure (27). The mean O-Si and Si-Si coordination numbers decrease in the presence of water: The hydrous values vary from 1.1 to 1.7 on compression (compared with anhydrous value of 1.4 to 2) and 1.8 to 4.7 (compared with anhydrous value of 2.5 to 5.5), respectively, over the compression range studied (17, 26).

The viscosity of silicate melts increases modestly along temperature profiles characteristic of Earth's interior because of the competing effects of pressure and temperature (Fig. 3). For example, along a slightly super-liquidus magma ocean isentrope (28), the viscosity of anhydrous melt increases by a factor of 10 from the surface to the core-mantle boundary. The variations are similar in size but non-monotonic along the estimated mantle solidus and liquidus (28). The non-monotonic variation of the viscosity along these curves is due to the rapid increase in temperature with increasing pressure at low pressure. The viscosity profiles of the hydrous melt show similar variations but are systematically shifted downward (Fig. 3).

Our results provide a fundamental basis for any dynamical model of magma ocean evolution. To illustrate, we use our viscosity results and those of previous ab initio simulations of silicate melts to estimate critical dynamical parameters. For a completely molten mantle such as one that may have occurred early in Earth's history (1), the estimated Rayleigh and Prandtl numbers (29) are $\sim 6 \times 10^{30}$ and ~ 60 , respectively, for the viscosity value of 0.048(10) Pa s for anhydrous $MgSiO_3$ liquid at mid-mantle condition (70 GPa

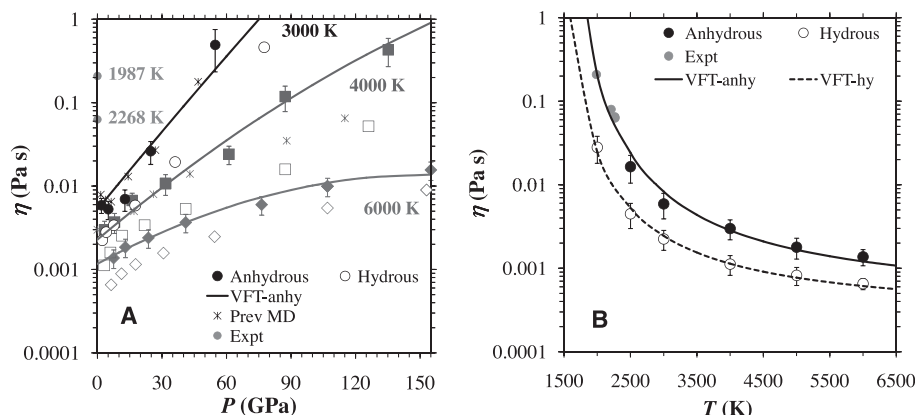
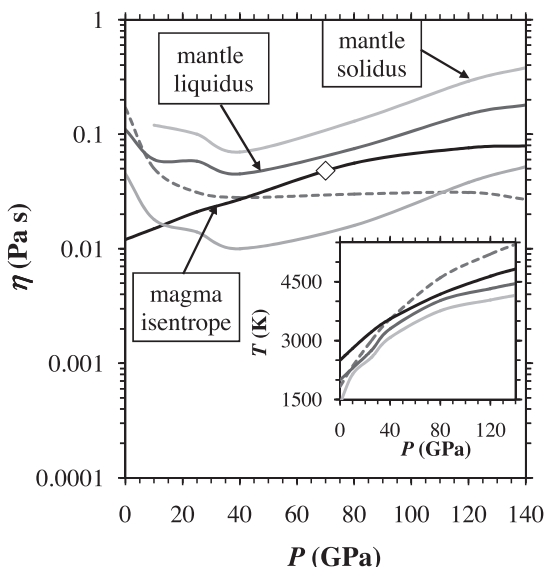


Fig. 2. Calculated viscosity (η) of anhydrous (anhy, solid symbols) and hydrous (hy, open symbols) $MgSiO_3$ melts. Our results are compared with experimental (Expt) data (6) at lower temperatures and ambient pressure for anhydrous $MgSiO_3$ liquid. Error bars indicate the statistical uncertainties. (A) Pressure variations along 3000 K (circles), 4000 K (squares), and 6000 K (diamonds) isotherms. The anhydrous results can be represented by the modified VFT (Vogel-Fulcher-Tammann) equation (35): $\eta(P, T) = \exp[-7.75 + 0.005 P - 0.00015 P^2 + (5000 + 135 P + 0.23 P^2)/(T - T_0)]$, with $T_0 = 1000$ K. Also shown are the results at 3000 K (14) and 4000 K (15) from previous molecular dynamics (MD) studies of anhydrous liquid based on semi-empirical pair potentials. (B) Temperature variations at the reference volume (V_r) together with experimental data (6) represented by the VFT equation, $\eta(T) = A \exp[B/(T - 1000)]$, where $A = 0.00033$ Pa s and $B = 6400$ K for anhydrous liquid and $A = 0.00024$ Pa s and $B = 4600$ K for hydrous liquid.

Fig. 3. Predicted viscosity for magma ocean based on anhydrous silicate liquid results (using the modified VFT relation shown in Fig. 2) along different temperature profiles (inset) from (28): magma ocean isentrope, mantle liquidus, mantle solidus, and melting curve of pure $MgSiO_3$ perovskite (dashed line). The hydrous result is shown only along the mantle liquidus (gray line). The symbol represents a viscosity value at mid-mantle condition.



and 4000 K along the magma ocean isentrope, Fig. 3). The Rayleigh number lies in the regime of turbulent convection: The presence of turbulence may substantially influence the settling of crystals as they form upon cooling. The surface heat flux, $F \sim 6 \times 10^4 \text{ W m}^{-2}$, estimated from mixing length theory far exceeds the incoming solar flux (30) and suggests that the surface temperature was set by heat exchange of the magma ocean with a dense silicate atmosphere rather than by solar radiation balance (2). This value of F implies a cooling time for the magma ocean $\sim 20 \text{ ky}$ (30). In fact, a number of processes are likely to increase the cooling time of the magma ocean substantially, including crystallization, which is predicted to initiate in the mid-mantle (28) and to separate the magma ocean into upper and basal layers (31). The evolution at this stage also depends strongly on the viscosity, which will set the time scale for buoyancy-driven motion of crystals and liquid that can lead to chemical differentiation. The direction of motion will be set by the crystal-liquid density contrast, the sign of which varies with pressure and temperature. Indeed, crystals are expected to float near the base of the mantle (16), producing a buoyantly stable basal magma layer that may be long-lived (31).

References and Notes

- R. M. Canup, *Icarus* **168**, 433 (2004).
- V. S. Solomatin, in *Evolution of the Earth*, D. Stevenson, Ed., vol. 9 of *Treatise on Geophysics*, G. Schubert, Ed. (Elsevier, Amsterdam, 2007), p. 91.
- S. E. Haggerty, V. Sautter, *Science* **248**, 993 (1990).
- J. Revenaugh, S. A. Sipkin, *Nature* **369**, 474 (1994).
- Q. Williams, E. J. Garnero, *Science* **273**, 1528 (1996).
- G. Urbain, Y. Bottinga, P. Richet, *Geochim. Cosmochim. Acta* **46**, 1061 (1982).
- I. Kushiro, H. S. Yoder Jr., B. O. Mysen, *J. Geophys. Res.* **81**, 6351 (1976).
- Y. Bottinga, P. Richet, *Geochim. Cosmochim. Acta* **59**, 2725 (1995).
- J. E. Reid et al., *Phys. Earth Planet. Inter.* **139**, 45 (2003).
- H. Behrens, F. Schulze, *Am. Mineral.* **88**, 1351 (2003).
- D. Tinker et al., *Am. Mineral.* **89**, 1701 (2004).
- C. Liebske et al., *Earth Planet. Sci. Lett.* **240**, 589 (2005).
- E. A. Wasserman, D. A. Yuen, J. R. Rustad, *Earth Planet. Sci. Lett.* **114**, 373 (1993).
- D. J. Lacks, D. Rear, J. A. Van Orman, *Geochim. Cosmochim. Acta* **71**, 1312 (2007).
- D. Nevins, F. J. Spera, M. S. Ghiorso, *Am. Mineral.* **94**, 975 (2009).
- L. Stixrude, B. B. Karki, *Science* **310**, 297 (2005).
- M. Mookherjee, L. Stixrude, B. B. Karki, *Nature* **452**, 983 (2008).
- Computations were performed by using the VASP software (32) with the local density approximation and ultrasoft pseudopotentials as before (16, 17, 26). Methods are available as supporting material on Science Online.
- L. A. Lange, *Rev. Mineral.* **30**, 331 (1994).
- M. L. Rivers, I. S. E. Carmichael, *J. Geophys. Res.* **92**, (B9), 9247 (1987).
- J. Mungall, *Geochim. Cosmochim. Acta* **66**, 125 (2002).
- T. K. Wan, T. S. Duffy, S. Scandolo, R. Car, *J. Geophys. Res.* **112** (B3), 03208 (2007).
- G. Giordano, D. B. Dingwell, *Earth Planet. Sci. Lett.* **208**, 337 (2003).
- C. A. Angell, P. A. Cheeseman, S. Tamaddon, *Science* **218**, 885 (1982).
- C. A. Angell, *Science* **267**, 1924 (1995).
- B. B. Karki, D. Bhattarai, M. Mookherjee, L. Stixrude, *Phys. Chem. Miner.* **37**, 103 (2010).
- E. M. Stolper, *Geochim. Cosmochim. Acta* **46**, 2609 (1982).
- L. Stixrude, N. de Koker, N. Sun, M. Mookherjee, B. B. Karki, *Earth Planet. Sci. Lett.* **278**, 226 (2009).

- The Rayleigh and Prandtl numbers are $Ra = [\alpha \rho g (T_M - T_S) L^3] / (\kappa \eta)$ and $Pr = \eta / (\rho \kappa)$. We have adopted the density ($\rho = 4410 \text{ kg m}^{-3}$) and thermal expansivity ($\alpha = 2.6 \times 10^{-5}$) (16) and assumed the depth scale $L = 3000 \text{ km}$, acceleration due to gravity $g = 10 \text{ m s}^{-2}$, thermal conductivity $k = 1.2 \text{ W m}^{-1} \text{ K}^{-1}$ (33), thermal diffusivity $\kappa = k / (\rho c_p)$ where the specific heat $c_p = 1660 \text{ J kg}^{-1} \text{ K}^{-1}$ (16), a mantle potential temperature $T_M = 2500 \text{ K}$, the lowest temperature at which the mantle will be completely molten, and a surface temperature $T_S = 1000 \text{ K}$ set by a dense atmosphere (2).
- The surface heat flux is $F = 0.22k(T_M - T_S)Ra^{2/7}Pr^{-1/7}L^{-1}$ (34) and the cooling time is $\tau_{cool} = T_M c_p M_M / 4\pi R^2 F$, where M_M is the mass of the mantle and R is the radius of Earth.
- S. Labrosse, J. W. Hernlund, N. Coltice, *Nature* **450**, 866 (2007).
- G. Kresse, J. Furthmüller, *Comput. Mater. Sci.* **6**, 15 (1996).
- A. M. Hofmeister, A. G. Whittington, M. Pertermann, *Contrib. Mineral. Petrol.* **158**, 381 (2009).
- B. I. Shraiman, E. D. Siggia, *Phys. Rev. A* **42**, 3650 (1990).
- K. R. Harris, L. A. Woolf, M. Kanakubo, *J. Chem. Eng. Data* **50**, 1777 (2005).
- This work was supported by NSF (EAR-0809489) and the UK National Environmental Research Council (NE/F01787/1). Computing facilities were provided by the Center of Computation and Technology at Louisiana State University. The authors thank J. Brodholt, M. Ghiorso, and S. Karato for useful comments and suggestions.

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/740/DC1
Materials and Methods
Figs. S1 and S2

16 February 2010; accepted 7 April 2010
10.1126/science.1188327

Extreme Deuterium Excesses in Ultracarbonaceous Micrometeorites from Central Antarctic Snow

J. Duprat,^{1*} E. Dobrică,¹ C. Engrand,¹ J. Aléon,¹ Y. Marrocchi,² S. Mostefaoui,² A. Meibom,² H. Leroux,³ J.-N. Rouzaud,⁴ M. Gounelle,² F. Robert²

Primitive interplanetary dust is expected to contain the earliest solar system components, including minerals and organic matter. We have recovered, from central Antarctic snow, ultracarbonaceous micrometeorites whose organic matter contains extreme deuterium (D) excesses (10 to 30 times terrestrial values), extending over hundreds of square micrometers. We identified crystalline minerals embedded in the micrometeorite organic matter, which suggests that this organic matter reservoir could have formed within the solar system itself rather than having direct interstellar heritage. The high D/H ratios, the high organic matter content, and the associated minerals favor an origin from the cold regions of the protoplanetary disk. The masses of the particles range from a few tenths of a microgram to a few micrograms, exceeding by more than an order of magnitude those of the dust fragments from comet 81P/Wild 2 returned by the Stardust mission.

The light element isotopic compositions of undifferentiated interplanetary material provide insights into the physicochemical processes that took place in the coldest regions of the early solar system. Large deuterium excesses are expected in the solid component(s) of comets because their water and HCN molecules exhibit D/H ratios from 2 to 15 times the terrestrial value, respectively (1). However, isotopic measurements

of fragments of comet 81P/Wild 2 returned by the Stardust mission show moderate D/H ratios that do not exceed three times the terrestrial value, possibly indicating a substantial alteration during impact capture process (2). By contrast, large D/H ratios have been observed as micrometer-sized hot spots in organic matter of interplanetary dust particles (IDPs) (3, 4) or primitive meteorites (5, 6). These D excesses may have been inherited from the cold

molecular cloud that predated the protosolar nebula (3, 5, 7) or may be the result of a local process that occurred in the cold outer regions of the protoplanetary disk (8, 9). The nature of the D-rich hot spots and their relationship to the organic matter bulk composition remain a matter of debate (6, 10). Organic matter in meteorites is sparse and disseminated in the matrix, with a maximum bulk concentration on the order of a few weight percent (wt %) (11). It is mainly accessible as the acid-insoluble component (IOM) remaining after demineralization of large amounts (grams) of primitive meteorites.

Large numbers of Antarctic micrometeorites (AMMs), which are IDPs with sizes ranging from 20 to 1000 μm , can be recovered from the Antarctic ice cap (12). Here, we describe AMMs obtained from the melting and sieving of 3 m^3 of ultraclean snow that fell in the vicinity of the French-Italian CONCORDIA station at Dome C

¹Centre de Spectrométrie Nucléaire et de Spectrométrie de Masse, Université Paris-Sud 11, CNRS/IN2P3, F-91405 Orsay, France. ²Laboratoire de Minéralogie et Cosmochimie du Muséum (LMCM), UMR 7202-CNRS INSU, Muséum National d'Histoire Naturelle, 57 Rue Cuvier, 75231 Paris Cedex 05, France. ³Unité Matériaux et Transformations, Université Lille 1 & CNRS, 59655 Villeneuve d'Ascq, France. ⁴Laboratoire de Géologie de l'École Normale Supérieure, UMR CNRS 8538, 24 rue Lhomond, 75231 Paris Cedex 5, France.

*To whom correspondence should be addressed. E-mail: jean.duprat@cnsn.in2p3.fr

(75°S, 123°E) between 1955 and 1970 (i.e., before human activities started in the area) (13). Because they were recovered from snow rather than ice, they did not endure the mechanical stress or aqueous alteration that are typical of both meteorites and AMMs recovered from Antarctic ice (14). Within the unmelted AMM population exhibiting a fine-grained, fluffy texture with no evidence for substantial heating during atmospheric entry (i.e., vesicles and/or magnetite shell), we report mineralogical and isotopic studies of two particles (particles 19 and 119) characterized by exceptionally high carbon content (Fig. 1A), referred to as ultracarbonaceous Antarctic micrometeorites (UCAMMs) (15). Their carbonaceous component is present in the form of organic matter and represents 48% and 85% of the analyzed area of particles 19 and 119, respectively (13). The sizes of particles 19 and 119 before fragmentation were 80 μm by 50 μm and 275 μm by 110 μm , respectively.

High-resolution transmission electron microscopy (HRTEM) of the UCAMMs shows a poor structural organization of the organic matter, with nanometer-sized polyaromatic layers, single or stacked in groups of two or three (Fig. 1B). This fringe length has the same scale as that reported in HRTEM studies of IOMs from Murchison and Orgueil carbonaceous chondrites (16). Combined studies of field emission gun scanning electron microscopy (FEG-SEM) and analytical TEM reveal mineral assemblages embedded within the organic matter (Fig. 1C) consisting mainly of Mg-rich silicates and Fe-Ni sulfides. These fine-grained mineral assemblages contain both amorphous and crystalline phases (Fig. 1C). Some glassy aggregates are comparable to the glass with embedded metal and sulfides (GEMS) identified in IDPs (Fig. 2) (13). The crystalline fraction of particle 19 is dominated by fine-grained clumps of olivines, pyroxenes, and iron sulfides, with mineral sizes ranging from ~15 to ~500 nm and having a pronounced frequency peak around 100 nm.

Isotopic imaging on fragments of particles 19 and 119 reveals large D excesses associated with the carbon-rich areas (Fig. 3) (13). By contrast, the carbon ($^{12}\text{C}/^{13}\text{C}$) and oxygen ($^{18}\text{O}/^{16}\text{O}$) isotopic maps do not exhibit large isotopic anomalies (13). The highest D/H ratio of $4.6 (\pm 0.5) \times 10^{-3}$ [$\delta\text{D} = 29,000 \pm 3000$ per mil (‰)] was measured within particle 119 (Fig. 3D). This ratio approaches that of the most D-rich hot spots observed by isotopic imaging in interplanetary material (3, 5). However, the D enrichments in the UCAMMs we analyzed have occurred over much wider areas than those reported previously in extraterrestrial materials. The areas with $\text{D}/\text{H} > 10^{-3}$ ($\delta\text{D} > 5400\text{‰}$) extend over 135 to 280 μm^2 and 65 to 200 μm^2 in particles 19 and 119, respectively (Fig. 3, A and C) (13). In the in situ isotopic imaging of IOM from primitive chondrites, the contribution of hot spots to the bulk D/H ratio remains small (5, 6), whereas in the UCAMMs the areas with D/H ratios larger than

10 times the terrestrial value represent most of the analyzed surface (Fig. 4) (13). Because the average D/H ratio of the UCAMMs is compatible with that of D-rich hot spots reported in other interplanetary materials, it is conceivable that these hot spots may result from the fragmentation and/or alteration of a primitive organic matter component similar to that revealed by this study.

In the D/H versus C/H atomic ratio plot, the data from particles 19 and 119 cluster along a main trend going from a first component at $\text{C}/\text{H} = 1$ to 2 and $\text{D}/\text{H} = 0.5 \times 10^{-3}$ to 1×10^{-3} up to a

high C/H component with C/H ratios ranging from 3 to 7 and D/H ratios ranging from 1×10^{-3} to 1.4×10^{-3} (Fig. 4) (13). The first component corresponds to that reported in bulk IOM analyses of CR carbonaceous chondrites (10). The nature of the UCAMM high C/H component remains to be clarified, but HRTEM data establish its highly disorganized structure and therefore rule out a substantial graphitization (Fig. 1B). The isotopic and elementary ratios of IOM from unequilibrated ordinary chondrites (UOCs) (10) plot along the same main trend

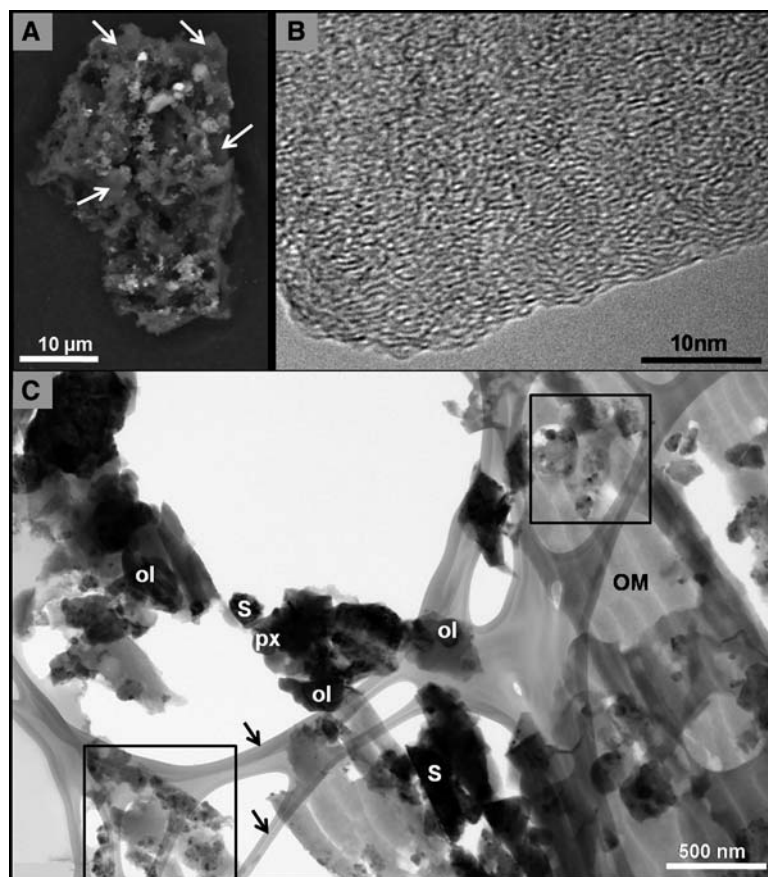
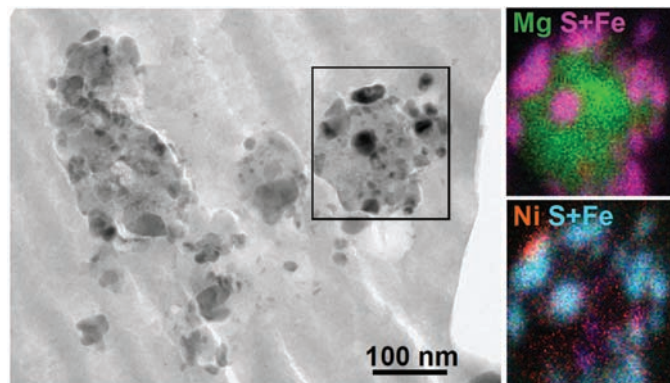


Fig. 1. (A) Backscattered scanning electron micrograph of particle 119. The carbon-rich areas appear dark (arrows); the bright inclusions are dominated by Fe-Ni sulfides and silicates. (B) High-resolution TEM image of particle 19. (C) Bright-field TEM image of particle 19. The lacey carbon film (13) is indicated as black arrows; the crystalline phases are Mg-rich olivines (ol), Mg-rich pyroxenes (px), and Fe-Ni sulfides (S); OM, organic matter. Glassy aggregates (GEMS candidates) are highlighted in black squares (13).

Fig. 2. (Left) Bright-field TEM image of a GEMS candidate embedded in the carbonaceous matter of particle 19. (Right) Stacked energy-dispersive x-ray spectroscopy elemental distribution maps of Mg, S + Fe (top) and Ni, S + Fe (bottom) for the GEMS candidate (black square).



(Fig. 4). It has been suggested that this main trend in UOCs may be due to a parent body process (10). However, it is difficult to conceive that the main trend from the UCAMMs may result from the same process, because (i) the UCAMMs and UOCs strongly differ both chemically and mineralogically [in particular, the UOC organic matter concentration (<0.5 wt %) (10)

is much lower than that observed in UCAMMs]; (ii) we do not observe any sign of thermal processing of the UCAMMs and, by contrast to the UOCs, we observe for $C/H > 3$ a constant D/H plateau rather than a correlated increase of D/H with C/H (fig. S2) (13); and (iii) the data presented here (Fig. 3) show that the whole range of variation of the main trend coexists within a few tens

of square micrometers, whereas the UOC main trend concerns IOM residues from the bulk meteorites. Therefore, the UCAMM main trend seems compatible with the sampling of a heterogeneous organic matter reservoir.

Above the main trend, the data broadly spread toward extreme D/H ratios with $D/H > 2.5 \times 10^{-3}$ and $C/H = 2$ to 6 (Fig. 4). The D -rich hot spots in IDPs for which C/H ratios have also been reported span a large range ($0 < C/H < 3$) (4), whereas the hot spots from IOMs of CR2 primitive chondrites have C/H ratios limited to a more restricted zone ($1 < C/H < 1.5$) (5). The UCAMM extreme D/H component seems to extend the high C/H trend observed in IDPs, including particles collected during the meteor shower associated with comet 26P/Grigg/Skjellerup (17).

High D excesses observed in interplanetary materials have long been attributed to interstellar chemistry, because large D enrichments ($D/H > 0.01$) are observed in the gas phase of cold molecular clouds (7). However, there is a strict upper limit on the fraction of crystalline relative to amorphous silicates in the interstellar medium (<0.2% by mass) (18). If the organic matter from the UCAMMs was a direct heritage of interstellar origin, one would expect the associated minerals to be dominated by amorphous silicates, which is not the case. Quite the opposite, the organic matter of the UCAMMs contains crystalline phases typical of silicates processed within the accretion disk (19), such as those observed both in anhydrous IDPs (20) and in the fine-grained fraction of Wild 2 particles (21). Therefore, the UCAMMs cannot be considered as a direct interstellar heritage but most probably sampled material (organic matter and minerals) from the protoplanetary disk itself.

Substantial D excesses have been identified at the molecular level in IOM from the Orgueil and Murchison meteorites, supporting an exchange mechanism between the organic matter and a local gaseous D -rich reservoir within the nascent solar system (9). Numerous astronomical observations demonstrate the occurrence of deuterated molecules in protoplanetary disks, some of them exhibiting large D/H variations ($0.01 < D/H < 0.1$) for radial distances between 30 and 70 AU (22). The large range of D/H ratios observed in the UCAMMs may be reminiscent of the D/H gradient that once existed at several tens of astronomical units from the young Sun.

Other than the bona fide Wild 2 particles returned by the Stardust mission, the assignment of a cometary or asteroidal origin to a given interplanetary dust particle remains speculative. The unmelted nature of the UCAMMs precludes high atmospheric entry velocities usually associated with a cometary origin. However, once released from their parent body, the trajectories of dust within that size range substantially evolve (as a result of resonances with giant planets, radiation pressure, the Poynting-Robertson effect, and solar wind drag), and cometary dust can

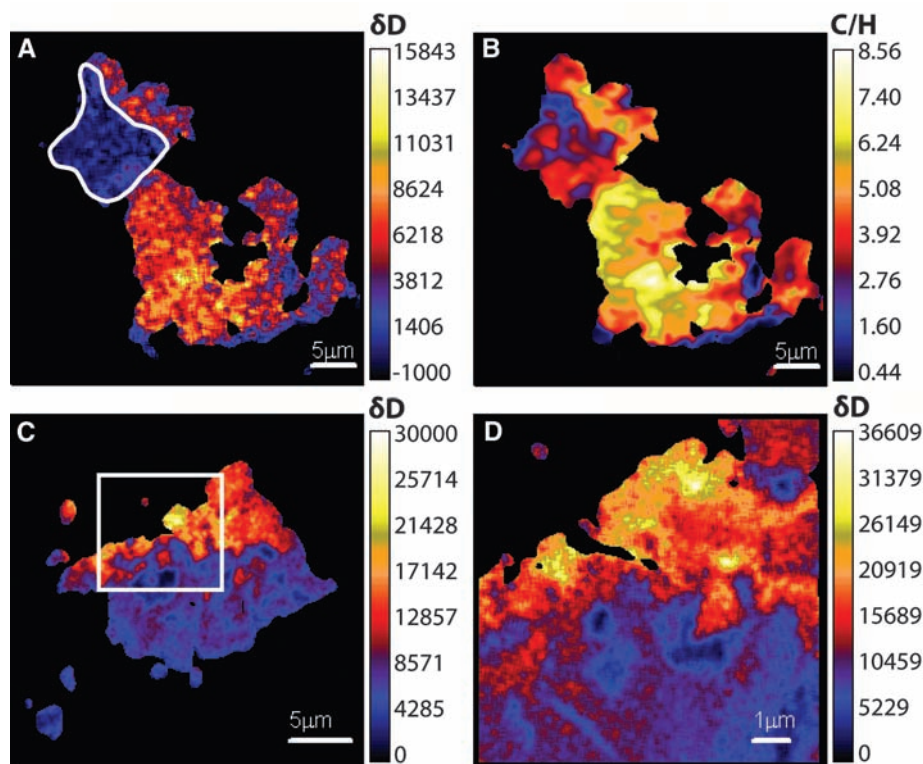


Fig. 3. NanoSIMS-50 (secondary ion mass spectrometry) isotopic and elemental maps of UCAMMs. (A and B) δD (‰) (13) (A) and C/H atomic ratio (B) of particle 19. The contour in (A) indicates a region with low D/H ratio (Fig. 4) (13). (C) δD (‰) map of particle 119. (D) Higher-magnification δD (‰) map of the zone indicated by the white rectangle in (C).

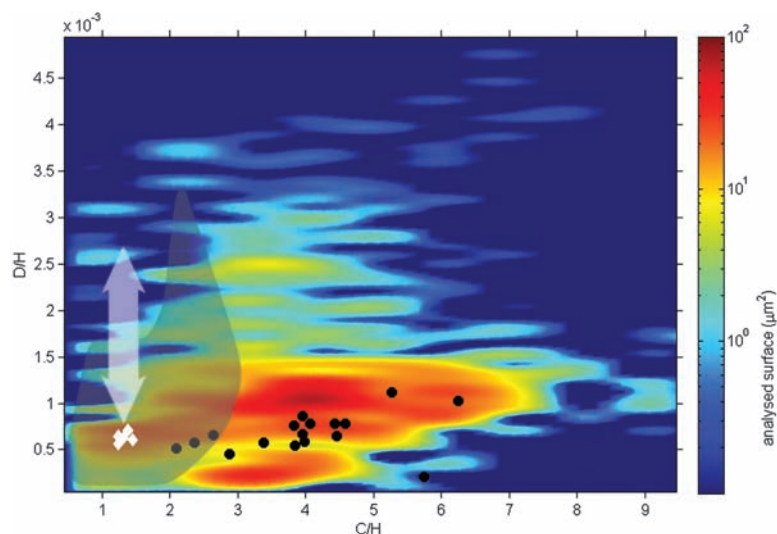


Fig. 4. Distribution of D/H versus C/H atomic ratios in particles 19 and 119 (13). The data from bulk IOM from CR (white diamonds), UOCs (black dots) (10), and the range of D -rich hot spots observed in primitive chondrites (5) (white arrow) and that from IDPs (4, 17) (gray surface) are reported.

eventually have Earth-crossing orbits that are difficult to distinguish from typical asteroidal trajectories (23). Some chondritic porous anhydrous IDPs are believed to be of cometary origin, as assessed by their noble gas content, high porosity, high carbon content, and high D/H ratios (3, 24, 25). The exceptionally high carbon content of the UCAMMs equals or exceeds that of the most C-rich IDPs (26, 27) and falls in the range of the CHON particles detected in comet 1P/Halley (28). Both the crystalline and amorphous silicates in the UCAMMs are comparable to those detected in the dust of different comets (29) and are compatible with the Wild 2 material (30). From these lines of evidence, we favor a cometary origin for these particles.

Both the large sizes and the high carbon content of the UCAMMs allow the direct in situ study of an early solar system organic compound together with its associated minerals, without any chemical extraction procedure. The association of extreme D-rich organic matter with high-temperature minerals confirms that material condensed or processed at close distances from the young Sun can be efficiently transported at several tens of astronomical units (21). The UCAMMs provide unique access to the intimate association of high- and low-temperature material from the protoplanetary disk in an unprecedented state of preservation.

References and Notes

1. R. Meier, T. C. Owen, *Space Sci. Rev.* **90**, 33 (1999).
2. K. D. McKeegan *et al.*, *Science* **314**, 1724 (2006).
3. S. Messenger, *Nature* **404**, 968 (2000).
4. J. Aléon, C. Engrand, F. Robert, M. Chaussidon, *Geochim. Cosmochim. Acta* **65**, 4399 (2001).
5. H. Busemann *et al.*, *Science* **312**, 727 (2006).
6. L. Remusat *et al.*, *Astrophys. J.* **698**, 2087 (2009).
7. T. J. Millar, A. Bennett, E. Herbst, *Astrophys. J.* **340**, 906 (1989).
8. Y. Aikawa, G. J. van Zadelhoff, E. F. van Dishoeck, E. Herbst, *Astron. Astrophys.* **386**, 622 (2002).
9. L. Remusat, F. Palhol, F. Robert, S. Derenne, C. France-Lanord, *Earth Planet. Sci. Lett.* **243**, 15 (2006).
10. C. M. O. D. Alexander, M. Fogel, H. Yabuta, G. D. Cody, *Geochim. Cosmochim. Acta* **71**, 4380 (2007).
11. I. Gilmour, in *Treatise on Geochemistry*, A. M. Davis, Ed. (Elsevier, Oxford, 2003), vol. 1, pp. 269–290.
12. M. Maurette *et al.*, *Nature* **351**, 44 (1991).
13. See supporting material on Science Online.
14. J. Duprat *et al.*, *Adv. Space Res.* **39**, 605 (2007).
15. T. Nakamura, T. Noguchi, Y. Ozono, T. Osawa, K. Nagao, *Meteorit. Planet. Sci.* **40** (suppl.), 5046 (2005).
16. S. Derenne, J. Rouzaud, C. Clinard, F. Robert, *Geochim. Cosmochim. Acta* **69**, 3911 (2005).
17. H. Busemann *et al.*, *Earth Planet. Sci. Lett.* **288**, 44 (2009).
18. F. Kemper, W. J. Vriend, A. G. G. M. Tielens, *Astrophys. J.* **609**, 826 (2004).
19. E. R. D. Scott, A. N. Krot, in *Chondrites and the Protoplanetary Disk*, A. N. Krot *et al.*, Eds. (Astronomical Society of the Pacific, San Francisco, 2005), pp. 15–53.
20. J. P. Bradley, *Geochim. Cosmochim. Acta* **58**, 2123 (1994).
21. D. Brownlee *et al.*, *Science* **314**, 1711 (2006).
22. C. Qi, D. J. Wilner, Y. Aikawa, G. A. Blake, M. R. Hogerheijde, *Astrophys. J.* **681**, 1396 (2008).
23. J. C. Liou, H. A. Zook, *Icarus* **123**, 491 (1996).
24. K. L. Thomas, L. P. Keller, G. E. Blanford, D. S. McKay, *Meteorit. Planet. Sci.* **27**, 296 (1992).
25. J. P. Bradley, in *Treatise on Geochemistry*, A. M. Davis, Ed. (Elsevier, Oxford, 2003), vol. 1, pp. 689–711.
26. K. L. Thomas, G. E. Blanford, L. P. Keller, W. Klöck, D. S. McKay, *Geochim. Cosmochim. Acta* **57**, 1551 (1993).
27. L. P. Keller *et al.*, *Geochim. Cosmochim. Acta* **68**, 2577 (2004).
28. M. E. Lawler, D. E. Brownlee, *Nature* **359**, 810 (1992).
29. D. H. Wooden, *Space Sci. Rev.* **138**, 75 (2008).
30. M. E. Zolensky *et al.*, *Science* **314**, 1735 (2006).
31. Supported by ANR grant 05-JC05-51407, FP6 Marie Curie Research Training Network "ORIGINS," ANR T-Tauri Chem, INSU (PNP), IN2P3, CNES, and CNRS. We are grateful to the French and Italian polar institutes, IPEV and PNRA, for their financial and logistic support. The NanoSIMS facility was established with funding from the CNRS, Région Île de France, MESR, and MNHN. The electron microscopy work in Lille is a national facility instrument INSU-CNRS. We are grateful to M. Maurette for stimulating exchanges and pioneering the micrometeorite research in Orsay; S. Derenne, E. Jacquet, E. Quirico, and G. Slodzian for enlightening discussions; and M. Horlok for careful reading of the manuscript.

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/742/DC1
Materials and Methods
Figs. S1 to S4
Table S1
References

17 November 2009; accepted 24 March 2010
10.1126/science.1184832

Cross-Reacting Antibodies Enhance Dengue Virus Infection in Humans

Wanwisa Dejnirattisai,¹ Amonrat Jumnainsong,¹ Naruthai Onsirisakul,^{1,2} Patricia Fitton,¹ Sirijitt Vasanawathana,³ Wanee Limpitikul,⁴ Chunya Puttikhunt,⁵ Carolyn Edwards,¹ Thaneeya Duangchinda,⁵ Sunpetchuda Supasa,⁶ Kriangkrai Chawansuntati,¹ Prida Malasit,^{5,6} Juthathip Mongkolsapaya,^{1,6*}† Gavin Screaton^{1*†}

Dengue virus co-circulates as four serotypes, and sequential infections with more than one serotype are common. One hypothesis for the increased severity seen in secondary infections is antibody-dependent enhancement (ADE) leading to increased replication in Fc receptor-bearing cells. In this study, we have generated a panel of human monoclonal antibodies to dengue virus. Antibodies to the structural precursor-membrane protein (prM) form a major component of the response. These antibodies are highly cross-reactive among the dengue virus serotypes and, even at high concentrations, do not neutralize infection but potently promote ADE. We propose that the partial cleavage of prM from the viral surface reduces the density of antigen available for viral neutralization, leaving dengue viruses susceptible to ADE by antibody to prM, a finding that has implications for future vaccine design.

Dengue virus (DENV) is a mosquito-borne virus infection found in tropical and subtropical areas of the world, with an estimated 50 to 100 million infections per year (1). A sequence variation of 30 to 35% allows DENV to be divided into four serotypes, and infection with one serotype does not provide protection from infection with the other serotypes, so that secondary or sequential infections are common (2, 3). Serious complications of dengue haemorrhagic fever (DHF) are more likely during secondary versus primary infections (2, 3).

In 1977, Halstead suggested antibody-dependent enhancement (ADE) to explain severe DENV infections (4). ADE has been widely studied and results from the high sequence divergence between DENV so that antibody to the first infection may not be of sufficient avidity to neutralize a secondary infection (5). The partial cross-reactivity may cause a degree of opsonization that promotes virus uptake into Fc-bearing cells such as monocytes and macrophages—a major site of DENV replication in vivo—leading to increased virus replication.

DENV envelope contains 180 copies of the E glycoprotein, which can be found in either dimeric or

trimeric conformation (6). The structural precursor-membrane protein (prM) is a 166-amino-acid protein intimately associated in a 1:1 fashion with domain II of E (7) and is believed to act as a chaperone for the folding of E and to prevent the premature fusion of virus to membranes inside the producing cell. prM contains a furin cleavage site and is cleaved into a C-terminal M portion containing a transmembrane domain that remains associated with the virus particle, and an N-terminal 91-amino-acid precursor fragment that dissociates upon release of the virus from the infected cell.

B cells from seven DENV-infected individuals (table S1) were used to produce human mAb by using the method of Traggiai (8). Culture supernatants were screened against structural antigens by using whole virus and against non-structural protein 1 (NS1) by means of enzyme-linked immunosorbent assay (ELISA).

Of 3020 cell lines, 301 screened positive, 73% reacted to the whole-virus ELISA for structural antigens, and 27% reacted to NS1. Positive super-

¹Department of Medicine, Imperial College London, London W12 0NN, UK. ²Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand. ³Pediatric Department, Khon Kaen Hospital, Khon Kaen 4000, Thailand. ⁴Pediatric Department, Songkhla Hospital, Songkhla 90100, Thailand. ⁵BIOTEC, NSTDA, Phatunthani 12120, Thailand. ⁶Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: g.screaton@imperial.ac.uk (G.S.); j.mongkolsapaya@imperial.ac.uk (J.M.)

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/745>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/328/5979/745/DC1>

This article **cites 15 articles**, 10 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/328/5979/745#otherarticles>

This article appears in the following **subject collections**:

Microbiology

<http://www.sciencemag.org/cgi/collection/microbio>

natants were tested for reactivity to specific DENV antigens by means of nonreducing Western blot (Fig. 1A). When the supernatants that reacted to whole DENV were tested, 78% gave a positive signal by means of Western blot, and all of these reacted to either E or prM, with no reactivity to capsid. The antibody-to-prM (anti-prM) response was substantial at 60% [95% confidence interval (CI), 67.3 to 52.2%] as compared with the response to E (40%); subgroup analysis of each of the individual cases is shown in table S2.

We next assessed the serotype specificity of the human antibodies by means of dot blot against the four viral serotypes, which showed a divergence in the cross-reactivity between the antibody-to-NS1 (anti-NS1) and structural [antibody-to-E (anti-E) and anti-prM] groups of antibodies. Half of the anti-NS1 showed limited cross-reactivity among DENV, whereas most of the antibodies directed to structural proteins showed full cross-reactivity against all virus serotypes (Fig. 1, B to D). Because these antibodies were made from secondary cases of DENV infection, we investigated primary anti-prM responses. Western blotting of DENV-infected cell lysates demonstrates that cross-reactive anti-prM responses are made during the primary infection (fig. S1), although as has been reported before, the anti-prM response is amplified after secondary infection (9).

Lastly, we tested cross-reactivity to the related flavivirus Japanese encephalitis virus (JEV), which co-circulates with DENV in some parts of south-east Asia (Fig. 1, E to G). Only 3% of the antibodies to prM cross-reacted with JEV, which is in contrast to the antibodies that recognized envelope, which showed 64% cross-reactivity. The relative specificity of anti-prM to DENV may reflect the lower sequence conservation between prM sequences (35% DENV versus JEV) as compared with that of E (50%); a comparison of sequence conservation among other members of the family flaviviridae can be found in table S3.

Six monoclonal anti-prM mAbs were produced; Western blotting showed that at least five of six react with the cleaved precursor peptide, and re-

activity was lost to reduced antigen, implying that they recognize conformational epitopes (fig. S2).

In general, the antibodies to prM were unable to completely neutralize infection (Fig. 2A). Instead, neutralization plateaued between 10 and 60%, and the partial neutralization was largely cross-reactive among the four virus serotypes; the only exceptions were mAb 5F9 and 135.3, both of which showed almost 100% neutralization of DENV4 at high antibody concentrations. This partial neutralization was in contrast to results seen with pooled convalescent dengue serum (PCS) or anti-envelope monoclonal antibodies, in which neutralization approached or reached 100%.

Next, we performed ADE assays using U937 cells as targets in which virus was preincubated with an increasing titer of antibody before addition to the Fc receptor-bearing cells. Enhancement of infection was seen with all six human monoclonal antibodies to prM with a peak of nearly a 10^5 -fold increase (Fig. 2B), which is consistent with a report of ADE with mouse anti-prM mAb (10).

To ascertain whether the results with these six anti-prM mAbs were representative, we tested the enhancing and neutralizing capacity of a further 20 anti-prM and 20 anti-E cell lines as well as two irrelevant human antibodies (Fig. 2, C and D). None of the antibodies to prM showed a high level of neutralization (19 of 19 showed <80%), whereas 12 of 20 antibodies to E showed >90% and 6 of 20 showed 100% neutralization. All of the antibodies to prM showed ADE of 10- to 800-fold, whereas the antibodies to E showed even more variable ADE (2- to 900-fold).

The failure of antibodies to prM to fully neutralize dengue viruses with a clear plateau in the response was puzzling and suggested that the virus may exist in two populations, one that is susceptible to neutralization and another that is not. Cleavage of prM during viral maturation is believed to be a prerequisite for viral replication, which is exemplified by the very low infectivity of DENV and tick-borne encephalitis

viral particles with wholly uncleaved prM (11–13). In many virus preparations, prM cleavage is incomplete, and cryogenic electron microscopy (cryo-EM) yields particles that contain both full-length prM and processed M protein, suggesting that a distribution of virus maturation may be present in virus cultures (14, 15). To our knowledge, it has not been formally demonstrated whether only fully processed virus is infectious or whether the virus can contain small numbers of prM molecules at its surface and still remain infectious. The demonstration here that the human antibodies to prM can show partial neutralization implies that some prM-containing particles remain infectious.

The propensity toward incomplete cleavage of prM in DENV leads to two interesting predictions. First, the density of prM at the surface of the virus may not be high enough to allow full neutralization with most antibodies to prM. Instead, viruses with low levels of prM may be susceptible to ADE. Second, viruses that are inherently non-infectious by virtue of displaying a high density of prM may be rendered infectious through ADE.

To investigate the effect of prM cleavage on neutralization and enhancement, virus was produced in cells cultured in the presence of ammonium chloride in order to raise intracellular pH and reduce the efficiency of furin cleavage (12). ELISA assays were performed in order to measure E and prM in the virus preparation; the E assay was calibrated by plotting a standard curve by use of recombinant E protein produced in Sf9 cells. A measure of the number of potential virus particles (virus-equivalent particles), which was derived assuming each particle contained 180 copies of E and the relative density of uncleaved prM, was expressed as the ratio of prM:E. The prM:E ratio was increased by roughly 40 and 80% when virus was cultured in 10 and 20 mM ammonium chloride, respectively (Fig. 3A). As expected, the infectivity of virus produced in the presence of NH_4Cl was markedly reduced from 46 to 555 virus-equivalent particles per focus-forming unit (FFU) (Fig. 3B). Although infectivity was reduced, infectious virus produced under

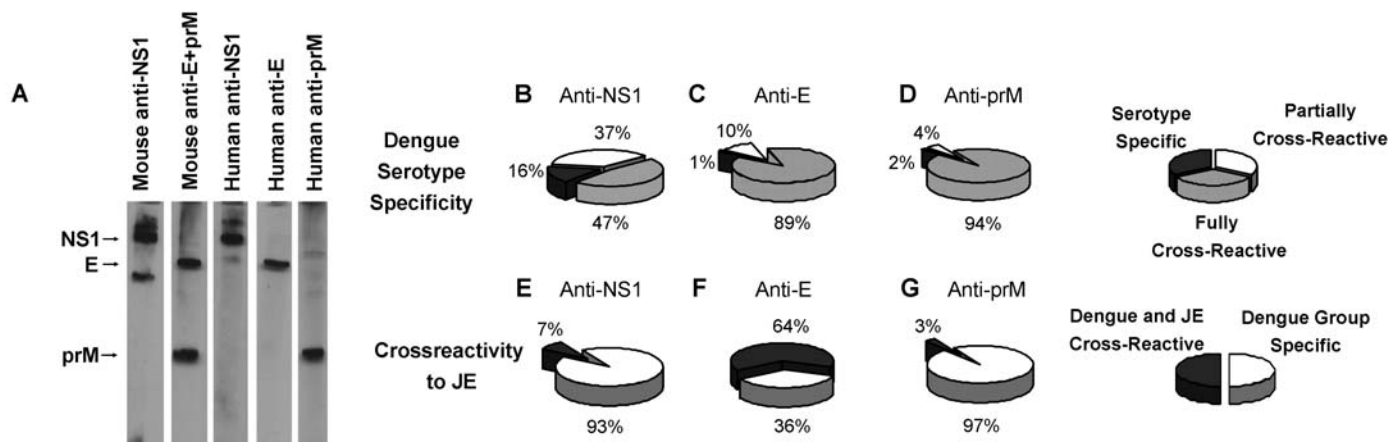


Fig. 1. Specificity of 301 human antibodies. (A) Western Blot of infected cell lysates (nonreduced, probed individually as single lane strips) showing reactivity of antibodies with dengue NS1, E, and prM proteins. (B to G) Cross-reactivity of human mAb within the DENV serotypes [(B) to (D)] or between the DENV group and JEV [(E) to (G)].

each condition remained partially susceptible to neutralization as before, and the titration curves for virus produced in 0, 10, and 20 mM NH_4Cl were similar (Fig. 3C).

We next tested enhancement of these viruses using either a constant amount of infectious virus—fixed FFUs (Fig. 3D)—or a constant amount of virus-equivalent particles (Fig. 3E). These results

show that the relatively poorly infectious virus cultured in the presence of NH_4Cl can be rendered much more infectious in the presence of enhancing antibodies to prM and indeed can be restored nearly to the level of control virus (Fig. 3E). These results were further exemplified by using virus that was produced in LoVo cells that lack functional furin and therefore produce virus with very low levels of cleaved prM (fig. S3, A to C) (13). Virus produced in LoVo cells as expected had a high prM:E ratio and very low infectivity ($< 10 \times 10^{-5}$ FFU per virus-equivalent particle) but infection could be enhanced in the presence of anti-prM.

Three populations of dengue virus appear to be produced: first, a population containing relatively high levels of prM that are inherently noninfectious but that can be made infectious in the presence of enhancing antibodies to prM; second, a population with an intermediate density of prM at the surface that can infect, but are susceptible to neutralization at high antibody titer; third, a population with low or absent prM at the surface that would not normally be susceptible to neutralization.

To test the relative roles of the antibodies to prM in neutralization and enhancement of primary cells, we looked at human monocytes, which are thought to be a major site of virus replication in vivo. Monocytes can be infected in the absence of antibody and, because they express Fc receptors, infection can be increased through ADE. To our surprise, human anti-prM mAbs failed to show any neutralization activity on primary monocytes and instead, even at concentrations of antibody as high as 30 $\mu\text{g}/\text{ml}$, enhanced infection from 20 to 70% (Fig. 4A) over a large range of antibody concentration.

Virus was generated in the insect cell line C6/36, which is known to cleave prM inefficiently, and the results we have obtained are therefore analogous to the first encounter with DENV, which was an insect-produced virus after a bite from an infected mosquito. Lastly, we set out to determine whether virus produced in primary mammalian cells contained noncleaved prM and whether anti-prM had any enhancing capacity on such virus. Virus was produced in immature dendritic cells (DCs) in which cleavage of prM was more efficient than in the insect cell culture but still not complete (Fig. 4B). As with insect-produced virus, the antibodies to prM were unable to fully neutralize DC-produced virus with a clear plateau in efficacy (Fig. 4C) but were still able to enhance infection, although to a lesser degree (Fig. 4D).

This is the first description, using human monoclonal antibodies, of the serological response in DENV infection. Anti-prM is a major component of the response, and most of these antibodies display limited virus neutralization capacity. The combination of partial cleavage of prM, together with substantial cross-reaction between serotypes, makes the anti-prM response particularly susceptible to enhancement. Promotion of such an antibody response toward anti-prM could thus be interpreted as immune evasion or even as an immune enhancement strategy of the virus.

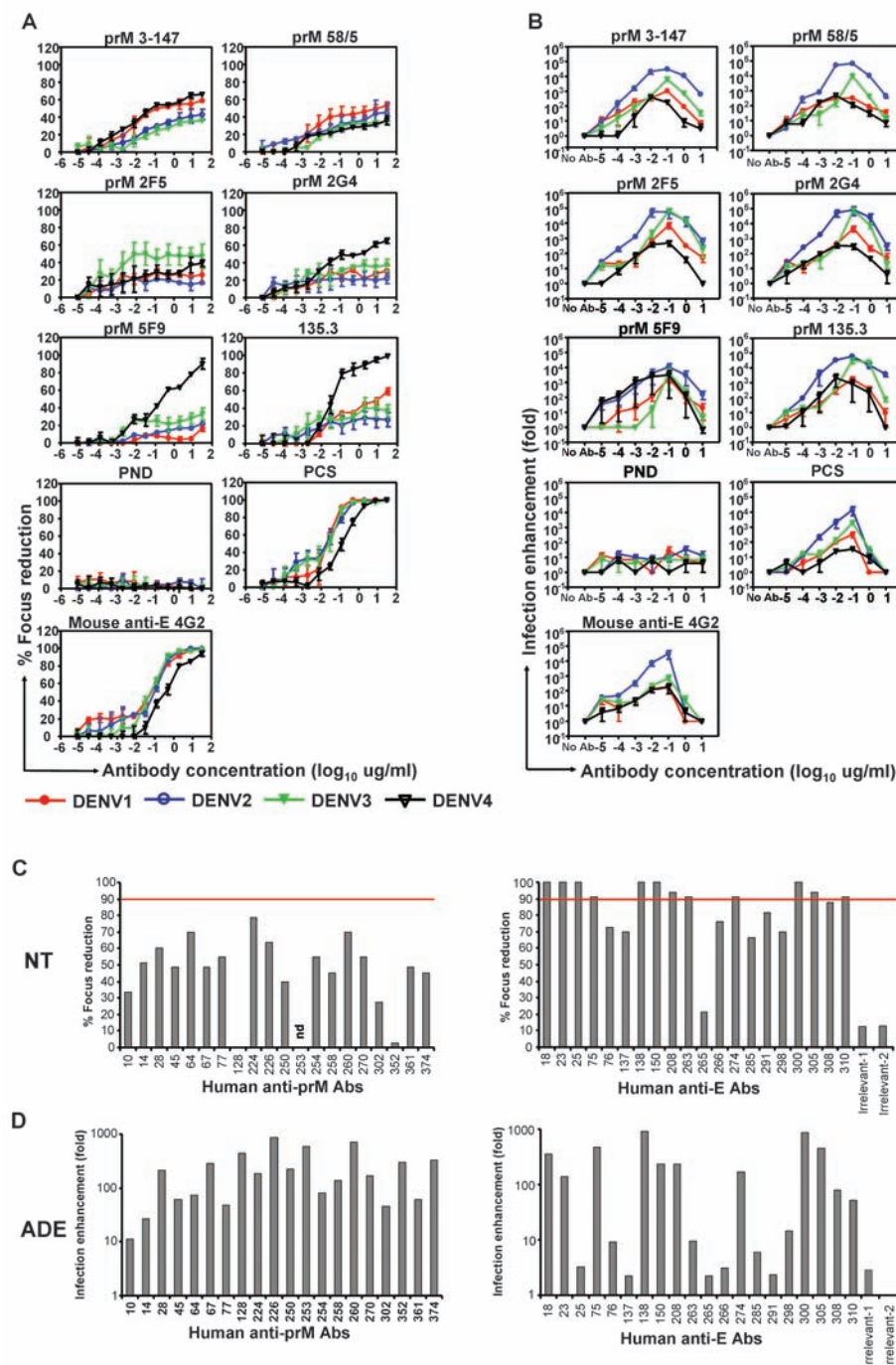
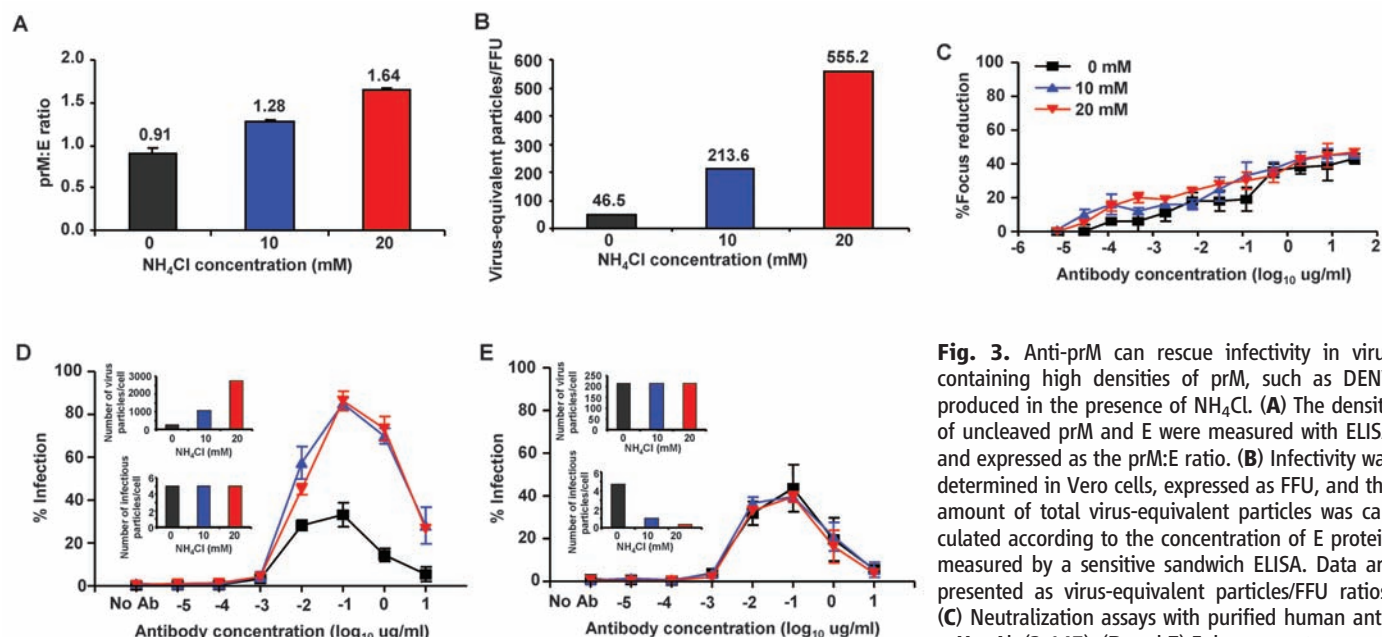


Fig. 2. Partial neutralization but potent enhancement by human monoclonal antibodies to prM. Neutralization assays (A) and enhancement assays (B) were performed with the six human anti-prM mAbs (clones 3-147, 58/5, 2F5, 2G4, 5F9, and 135.3), mouse anti-E mAb (4G2), and purified immunoglobulin (Ig) from pooled dengue convalescent serum (PCS) and pooled non-dengue immune serum (PND) were used as controls (mean \pm SE from three independent experiments). Culture supernatants from 20 anti-E and 20 anti-prM cell lines, all of which were specific to DENV2 and cross-reactive with other DENV serotypes, were assayed in neutralization (C) and ADE assays (D) to DENV2 strain 16681. Neutralization was performed by means of focus-forming assay on Vero cells by using a 1:2 dilution of supernatant, whereas ADE was performed by using a 1:100 dilution on U937 cells; infection was read with fluorescence-activated cell sorting (FACS) by using 4G2.



U937 cells read out by FACS based intracellular staining for DENV antigens (4G2) using either (D) a constant amount of infectious virus or (E) constant number of virus particles (mean \pm SE from three independent experiments).

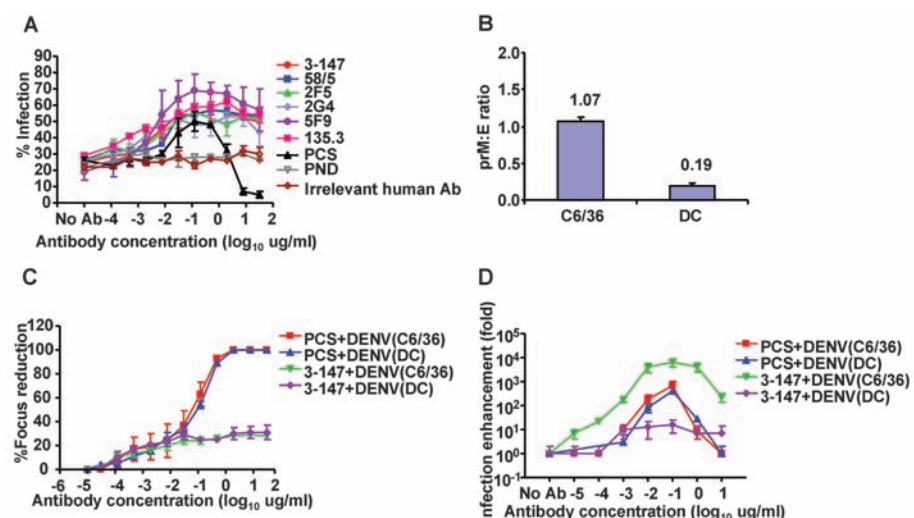


Fig. 4. The roles of anti-prM on neutralization and enhancement of DENV infection of peripheral blood mononuclear cells (PBMCs). (A) PBMCs were infected with DENV2 in the presence of human anti-prM mAbs; at 24 hours, DENV Ag was stained intracellularly (4G2) and detected with flow cytometry in gated monocytes. PCS, PND, and irrelevant human mAb were used as control. (B) The density of prM on DENV from C6/36 cells and DC were detected with ELISA and presented as prM:E ratio. (C) Neutralization and (D) ADE of infection performed on Vero and U937 cells, respectively, of DENV generated from either C6/36 cells or DC in the presence of PCS or anti-prM mAb (3-147) (mean \pm SE from three independent experiments).

Most current DENV vaccine candidates—whether naturally attenuated, recombinantly attenuated, or chemically inactivated virus or DENV–yellow fever chimeras—contain native dengue prM sequences (16). It may be advisable to design DENV vaccines that minimize the anti-prM response. There is relatively low sequence conservation between DENV prM and sequences from other flaviviruses, and the majority of antibodies to dengue prM do not cross-react with JEV. Chimeric attenuated viruses containing

heterologous flaviviral prM sequences may therefore not lead to such cross-reactive anti-prM responses as seen on infection with viruses that contain native dengue prM sequences, although the feasibility of making such chimeras has to our knowledge not yet been tested.

References and Notes

- World Health Organization, Fact sheet no. 117 (2008); available at www.who.int/mediacentre/factsheets/fs117/en/.
- M. G. Guzmán et al., *Am. J. Epidemiol.* **152**, 793, discussion 804 (2000).

Fig. 3. Anti-prM can rescue infectivity in virus containing high densities of prM, such as DENV produced in the presence of NH_4Cl . (A) The density of uncleaved prM and E were measured with ELISA and expressed as the prM:E ratio. (B) Infectivity was determined in Vero cells, expressed as FFU, and the amount of total virus-equivalent particles was calculated according to the concentration of E protein measured by a sensitive sandwich ELISA. Data are presented as virus-equivalent particles/FFU ratios. (C) Neutralization assays with purified human anti-prM mAb (3-147). (D and E) Enhancement assays of

- N. Sangkawibha et al., *Am. J. Epidemiol.* **120**, 653 (1984).
- S. B. Halstead, E. J. O'Rourke, *J. Exp. Med.* **146**, 201 (1977).
- M. S. Diamond, T. C. Pierson, D. H. Fremont, *Immunol. Rev.* **225**, 212 (2008).
- Y. Modis, S. Ogata, D. Clements, S. C. Harrison, *Nature* **427**, 313 (2004).
- L. Li et al., *Science* **319**, 1830 (2008).
- E. Traggiai et al., *Nat. Med.* **10**, 871 (2004).
- C.-Y. Lai et al., *J. Virol.* **82**, 6631 (2008).
- K.-J. Huang et al., *J. Immunol.* **176**, 2825 (2006).
- F. Guirakhoo, F. X. Heinz, C. W. Mandl, H. Holzmann, C. Kunz, *J. Gen. Virol.* **72**, 1323 (1991).
- I. M. Yu et al., *Science* **319**, 1834 (2008).
- I. A. Zyburt, H. van der Ende-Metselaar, J. Wilschut, J. M. Smit, *J. Gen. Virol.* **89**, 3047 (2008).
- J. Junjhon et al., *J. Virol.* **82**, 10776 (2008).
- M. V. Cherrier et al., *EMBO J.* **28**, 3269 (2009).
- S. S. Whitehead, J. E. Blaney, A. P. Durbin, B. R. Murphy, *Nat. Rev. Microbiol.* **5**, 518 (2007).
- We thank W. Supanchaimat, V. Jarupoonphol, S. Jinathongthai, K. Sriurksa, P. Wongsilarat, T. Suphachaiyakit, K. Ratarpa, Y. Sutvigat, and the staff of Khon Kaen and Songkhla hospitals for sample collection; L. Damrikarnlerd, P. Suriyapol, C. Komoltri, S. Udompunturak, N. Tangthawornchaikul, A. Jirangsri, K. Sae-Jang, and S. Supajitkasem for data and clinical database management and statistical analysis; N. Sittisombut and P. Keelapang from Chiang Mai University for sharing knowledge; and C. P. Simmons and B. Wills from the Oxford University Clinical Research Unit, Ho Chi Minh City for samples. This work was supported by the Medical Research Council, UK; the Wellcome Trust, UK; the National Institute for Health Research Biomedical Research Centre funding scheme; the Thailand Tropical Disease Research Program T2; and the Thailand National Centre for Genetic Engineering and Biotechnology.

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/745/DC1
Materials and Methods
Figs. S1 to S3
Tables S1 to S4

25 November 2009; accepted 22 March 2010
10.1126/science.1185181

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/749>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/science.1185837/DC1>

This article **cites 27 articles**, 7 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/328/5979/749#otherarticles>

This article has been **cited by** 1 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/cgi/content/full/328/5979/749#otherarticles>

This article appears in the following **subject collections**:

Immunology

<http://www.sciencemag.org/cgi/collection/immunology>

Induction of Lymphoidlike Stroma and Immune Escape by Tumors That Express the Chemokine CCL21

Jacqueline D. Shields,* Iraklis C. Kourtis,* Alice A. Tomei, Joanna M. Roberts, Melody A. Swartz†

Tumor manipulation of host immunity is important for tumor survival and invasion. Many cancers secrete CCL21, a chemoattractant for various leukocytes and lymphoid tissue inducer cells, which drive lymphoid neogenesis. CCL21 expression by melanoma tumors in mice was associated with an immunotolerant microenvironment, which included the induction of lymphoid-like reticular stromal networks, an altered cytokine milieu, and the recruitment of regulatory leukocyte populations. In contrast, CCL21-deficient tumors induced antigen-specific immunity. CCL21-mediated immune tolerance was dependent on host rather than tumor expression of the CCL21 receptor, CCR7, and could protect distant, coimplanted CCL21-deficient tumors and even nonsyngeneic allografts from rejection. We suggest that by altering the tumor microenvironment, CCL21-secreting tumors shift the host immune response from immunogenic to tolerogenic, which facilitates tumor progression.

Cancer fate, including progression, metastasis, and therapy resistance, is largely determined by the interactions between a tumor and host immune cells. Immune cells can recognize tumors by their antigenic profiles, but many tumors manipulate these cells to escape immune surveillance. To accomplish this, tumors can mimic immune signaling pathways that alter the tumor microenvironment to favor the activation of regulatory T (T_{reg}) cells and suppress

effector functions (1–3), driving immunological tolerance and tumor progression.

Here, we examine a mechanism of tumor-induced immune tolerance that bears similarities to the tolerance-maintaining functions of the lymph node (LN) stroma. In the lymph node paracortex, specialized stromal cells called fibroblastic reticular cells (FRCs) secrete the CCR7 ligands CCL21 and CCL19, which guide the interactions between CCR7⁺ T cells and

antigen-presenting cells (APCs) needed for T cell education and priming. Although these events are sufficient to trigger adaptive immunity, they are also necessary for maintaining peripheral tolerance, because T_{reg} cells require LN occupancy and CCR7 signaling for their activation and function (4–6), and the loss of CCR7 signaling is associated with spontaneous autoimmunity (7, 8). The lymph node stroma itself can also promote deletion of self-reactive cells (9, 10) and can help to maintain homeostasis of naïve T cells (11). We recently showed that invasive tumor cells secrete CCL21 (12), and we verified this here in several invasive human tumor lines cultured in three-dimensional conditions (fig. S1A). Given the critical role of CCR7 in both immunity and tolerance, we asked how endogenous tumor CCL21 expression would affect the host immune response.

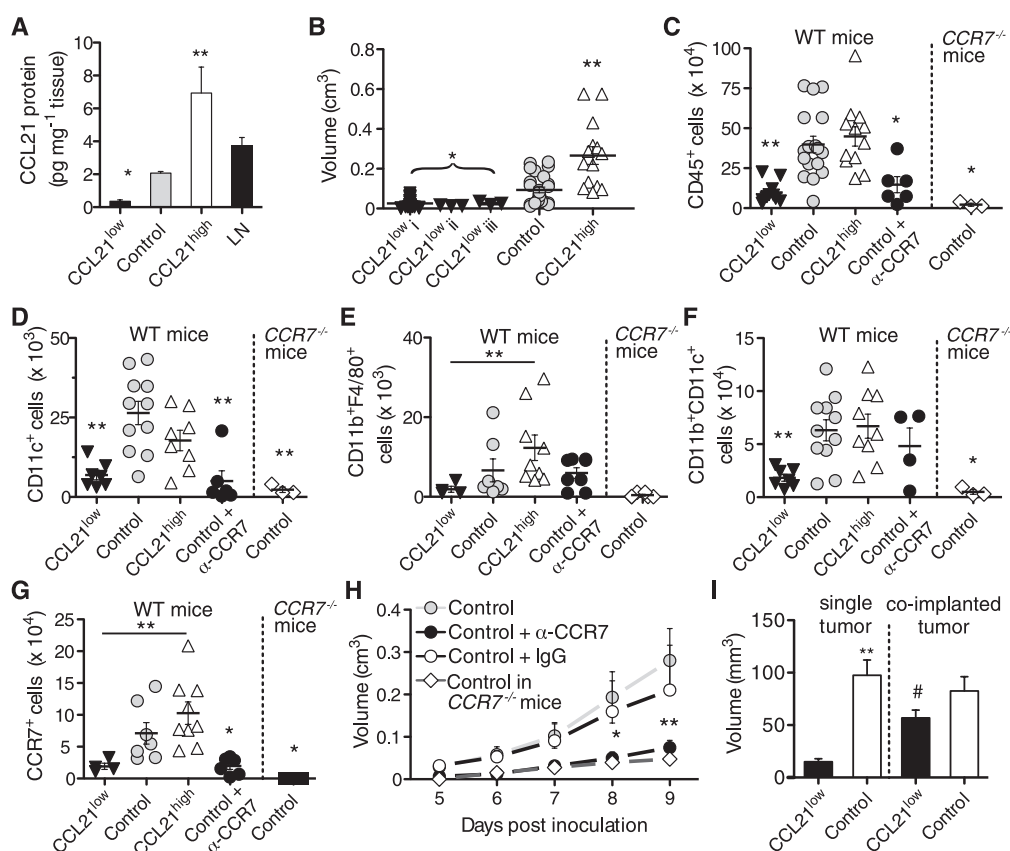
Aside from recruiting leukocytes and guiding their interactions in the LN, CCL21 is also a main driver of lymphoid tissue formation (13–15), as it attracts CCR7⁺ lymphoid tissue inducer (LTi) cells that drive the maturation of lymphoid stroma (16). Notably, expression of CCL19 and CCL21 in nonlymphoid tissues has been correlated with autoimmunity and inflammation, as

Institute of Bioengineering, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland.

*These authors contributed equally.

†To whom all correspondence should be addressed. E-mail: melody.swartz@epfl.ch

Fig. 1. CCL21 expression promotes tumor growth that is host CCR7-dependent. **(A)** CCL21 protein levels (determined by ELISA) in tumor and naïve lymph node lysates 9 days post-implantation (p.i., $n \geq 4$). **(B)** Day 9 p.i. tumor volumes. Multiple CCL21^{low} clones were implanted (CCL21^{low}i, clone 21/217; CCL21^{low}ii, clone 21/217 D8; CCL21^{low}iii, clone 21/401 H5; all $n \geq 3$; control, $n = 22$; CCL21^{high}, $n = 10$). Bars show medians \pm SE. **(C)** Tumor infiltration of CD45⁺ leukocytes as determined by flow cytometry, and **(D to F)** APC subpopulations at day 9 p.i. in wild-type ($n \geq 10$) and CCR7^{-/-} mice ($n = 3$); bars show medians \pm SE. **(G)** CCR7⁺ leukocyte infiltrates in tumors 9 days p.i. in wild-type ($n \geq 4$) and CCR7^{-/-} mice ($n = 3$). Bars show medians \pm SE. **(H)** Day 9 p.i. control tumor volumes in wild-type mice treated with CCR7 neutralizing antibodies or control IgG, or in CCR7^{-/-} mice ($n \geq 3$). Data represent means \pm SEM. **(I)** Volumes of single ($n = 22$) and co-implanted tumors ($n \geq 7$) at day 9 p.i. Data represent means \pm SEM. * $P < 0.05$, ** $P < 0.01$ relative to control tumors, and # $P < 0.05$ relative to single implanted CCL21^{low} tumors, one-way analysis of variance (ANOVA) and Bonferroni post-test adjustment.



well as immune suppression (7, 15, 17). Likewise, exogenous CCR7 ligands have been demonstrated to induce both antitumor immunity and tumor immune suppression (7, 18, 19). Here, in contrast to the studies using exogenous CCR7 ligands, we examine the effects of endogenous melanoma CCL21 expression on tumor fate.

We engineered three stable cell sublines derived from murine B16-F10 melanomas to knockdown endogenous CCL21 secretion by shRNA (CCL21^{low}), to express endogenous amounts (scrambled shRNA control) at amounts comparable to those measured in normal LNs, or to overexpress CCL21 (CCL21^{high}) (20) (Fig. 1A). Surprisingly, when implanted into immune competent syngeneic C57BL/6 mice, CCL21^{high} and control tumors grew significantly larger than CCL21^{low} tumor clones (Fig. 1B)—even though the CCL21-expressing tumors attracted more CD45⁺ leukocytes, including APCs and CCR7⁺ leukocytes (Fig. 1, C to G, and fig. S1, B to E). Note that CCR7⁺ APCs in CCL21-expressing tumors retained the capacity to traffic from the tumor to the draining LNs upon uptake of 0.5 μ m fluorescent beads (fig. S2), which suggested that their ability to uptake antigen was not impaired.

These differences in growth could have resulted either from the host response to the tumor or from autocrine effects of CCL21 signaling on tumor cells themselves, because they express CCR7 (fig. S1F). In vitro, however, the three different cell lines proliferated, formed spheroids, and migrated up a gradient of exogenous CCL21 similarly (fig. S1, G to J); furthermore, these behaviors were unaltered by the addition of exogenous CCL21 protein or with CCR7-blocking antibodies, which indicated that the differences seen in growth were dependent on the in vivo environment.

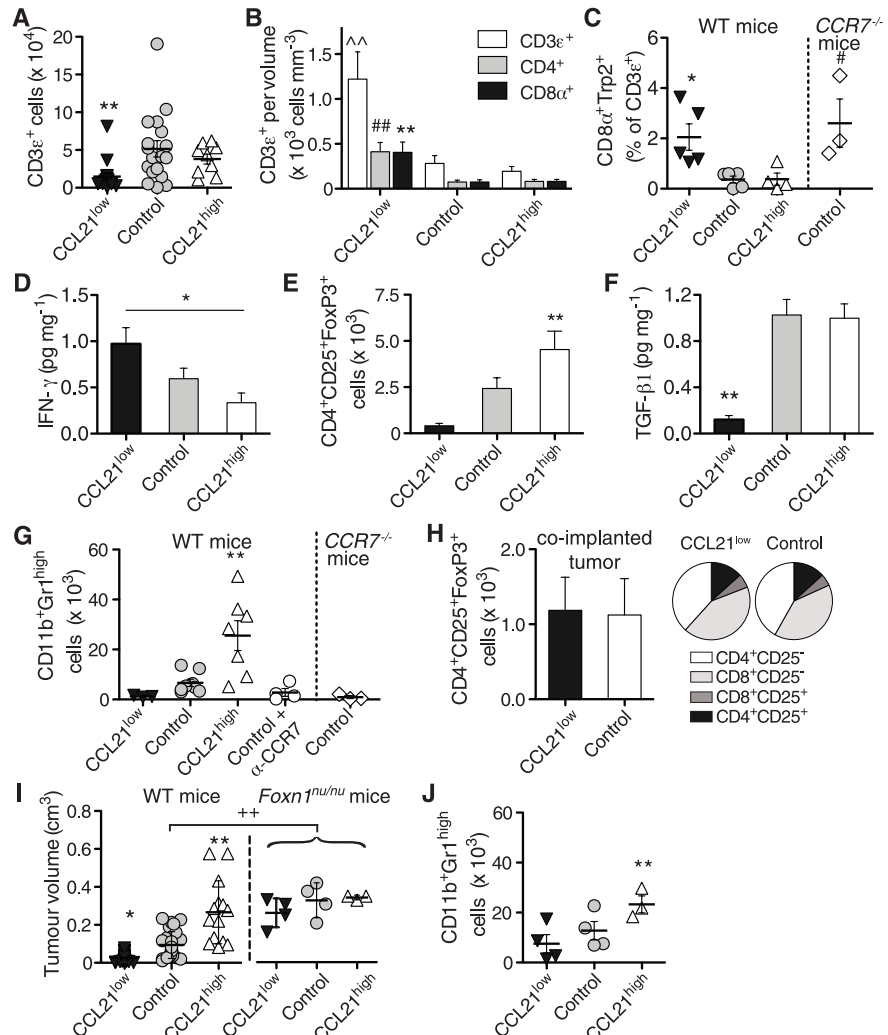
Moreover, tumor growth was host CCR7-dependent, because control tumors grew poorly when implanted into CCR7-deficient mice or into wild-type mice treated systemically with CCR7-blocking antibodies (Fig. 1, C to H). Therefore, tumor-mediated, CCL21-dependent modulation of the host response, rather than autocrine effects on the tumor itself, was responsible for the differential tumor propagation observed.

To further verify that these growth differences were due to variations in the host immune response to the tumor, rather than to changes incurred on the tumor cells themselves by autocrine CCR7 signaling, we coimplanted CCL21^{low} and

control tumors into the same mouse, each on opposite shoulders. In these mice, control tumors could rescue the growth of CCL21^{low} tumors to control levels (Fig. 1I), which further demonstrated that a host response was responsible for the differences in tumor growth seen earlier.

Given this apparent contradiction—increased tumor growth associated with enhanced leukocyte attraction and normal APC trafficking to LNs—we next asked how the secretion of CCL21 could affect interactions between the tumor and its immune cell infiltrates. Upon examining the T cell populations within the tumors, we found that, although control and CCL21^{high} tumors attracted more T cells overall (Fig. 2A), CCL21^{low} tumors contained higher densities of T cells (Fig. 2B) and more melanoma antigen [tyrosinase-related protein 2 (Trp2)]-specific CD8⁺ T cells (Fig. 2C). This was consistent with increased amounts of interferon- γ (IFN- γ), interleukin 2 (IL-2), and IL-4 in CCL21^{low} tumors (Fig. 2D and fig. S4, A and B), cytokines that are all associated with cytotoxic T cell responses and antitumor immunity (1–3). In contrast, control and CCL21^{high} tumors contained more CD4⁺ CD25⁺ FoxP3⁺ T_{reg} cells (Fig. 2E and fig. S3, A to C) and higher amounts of transforming growth

Fig. 2. CCL21 expression leads to a tolerogenic tumor microenvironment. **(A)** Total CD3 ϵ^+ T cells within CCL21^{low}, control, and CCL21^{high} ($n = 9$) tumors ($n \geq 9$). Bars show medians \pm SE. **(B)** Number of T cells per unit tumor volume ($n \geq 9$). Data represent means \pm SEM. **(C)** Frequency of tyrosinase-related protein-2 peptide SVYDFVWL (Trp2₁₈₀₋₁₈₈)-specific T cells (CD19⁺CD3 ϵ^+ CD8 α^+ SVYDFVWL-MHC pentamer⁺) (29) within tumors of wild-type and CCR7^{-/-} mice, as determined by flow cytometry. Bars show medians \pm SE. **(D)** IFN- γ protein levels within tumors as determined by enzyme-linked immunosorbent assay (ELISA) ($n = 8$). Data represent means \pm SEM. **(E)** Quantification of tumor-infiltrated T_{reg} cells within CCL21^{low}, control, and CCL21^{high} tumors ($n \geq 5$). Data represent means \pm SEM. **(F)** Total TGF- β 1 protein levels within tumors as determined by ELISA ($n \geq 4$). Data represent means \pm SEM. **(G)** CD11c⁺CD11b⁺F4/80⁺Gr1^{high} myeloid-derived suppressor tumor infiltrates ($n \geq 6$). Bars show medians \pm SE. **(H)** Intratumoral T cell populations and T_{reg} cells within coimplanted CCL21^{low} and control tumors ($n \geq 4$). Data represent means \pm SEM. **(I)** Comparison of tumor volumes in wild-type ($n \geq 10$) and athymic mice (Foxn1^{nu/nu}, $n = 4$). Bars show medians \pm SE. **(J)** CD11c⁺CD11b⁺F4/80⁺Gr1^{high} myeloid-derived suppressor tumor infiltrates in Foxn1^{nu/nu} mice ($n = 4$). Bars show medians \pm SE. All data were taken at day 9 p.i. The symbols * or # indicate $P < 0.05$, and double superscripts indicate $P < 0.01$ compared with respective control tumors.



factor- β 1 (TGF- β 1) (Fig. 2F). TGF- β 1 is a key regulator of tumor tolerance that suppresses antigen-specific CD8⁺ T cell function, promotes T_{reg} cell induction, and shifts the macrophage populations from classically activated (M1) to alternatively activated, protumor (M2) phenotypes (21, 22). Furthermore, tumor expression of CCL21 led to enhanced CCR7-dependent attraction of CD11b⁺CD11c⁺F4/80⁺Gr1^{high} myeloid-derived suppressor cells (MDSCs) that also expressed inducible nitric oxide synthase (Fig. 2G and fig. S4G), a cell type known to drive tumor progression (1–3). Coincident with more MDSCs were higher amounts of the MDSC chemoattractants CCL2 and C5a in CCL21-expressing tumors (fig. S4, D and E). Finally, when implanted into the same mouse but on opposite shoulders, CCL21^{low} and CCL21^{high} tumors (which grew similarly large) (Fig. 1I) displayed similar distributions of T cell populations (Fig. 2H), in-

cluding T_{reg} cell numbers similar to those seen in CCL21^{high} tumors grown alone.

The central role of the adaptive immune response in the prevention of CCL21^{low} tumor establishment was further demonstrated by using athymic *Foxn1*^{nu/nu} mice, which lack T cells and thus cell-mediated immunity. In these mice, CCL21^{low} tumors grew just as large as control and CCL21^{high} tumors (Fig. 2I), despite their impaired recruitment of MDSCs (Fig. 2J).

We next examined the tumor stroma, because peripheral expression of CCL21 can drive lymphoid neogenesis via recruitment of CCR7⁺ LTi cells (10, 13–15, 23). In the LN paracortex, FRCs are the major source of CCR7 ligands and are characterized by glycoprotein gp38 and the ER-TR7 antigen. In control and CCL21^{high} tumors, we observed FRC networks at the tumor margins that were reminiscent of those in the LN paracortex (Fig. 3, A and B, and fig. S5). These

tumor stromal networks expressed gp38, which in tumors has been associated with poor prognosis (24). CCL21-expressing tumors, but not CCL21^{low} tumors, also expressed the catabolic enzyme indoleamine 2,3-dioxygenase (IDO) (Fig. 3C), a potent tumor immune suppressor (25), as well as complement receptor 1–related gene/protein γ (Crry) (Fig. 3D), a complement-regulating protein that helps maintain self-tolerance and that can inhibit antitumor immunity (26). Blood vessel density appeared similar in all tumors (Fig. 3E), but some vessels in control and CCL21^{high} tumors also expressed peripheral node addressin (PNAd), which is normally associated with LN high endothelial venules (HEVs) (fig. S5C).

Consistent with these LN-like stromal changes, CD45⁺CD3 ϵ ⁺CD4⁺ROR γ t⁺ LTi cells were preferentially recruited to control and CCL21^{high} tumors in a host CCR7-dependent manner,

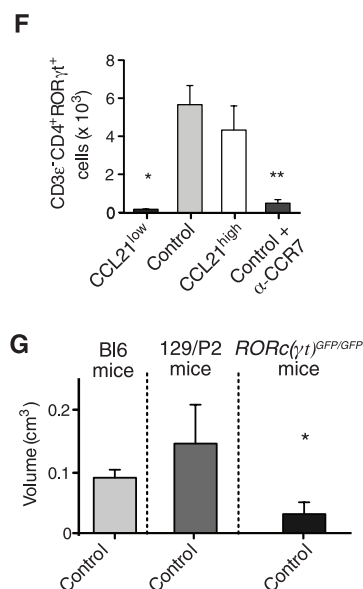
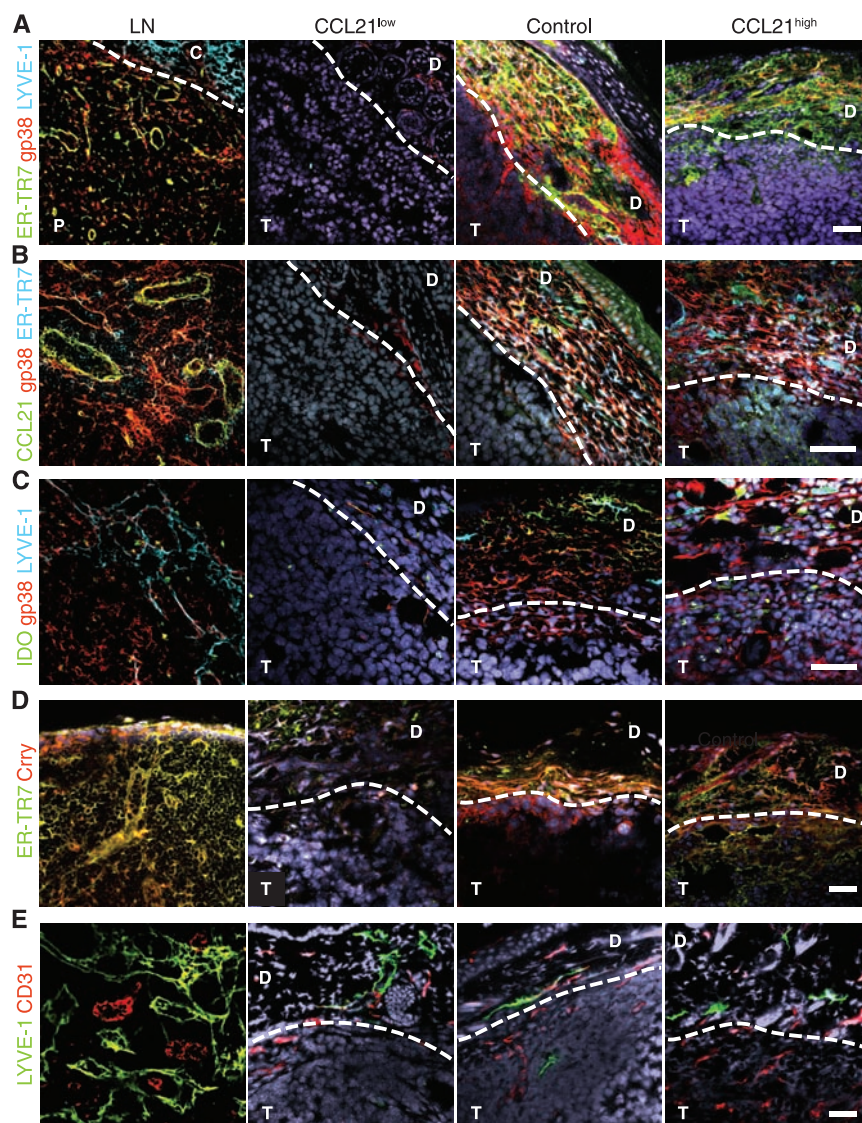


Fig. 3. CCL21-expressing tumors develop stromal zones reminiscent of lymph node paracortex stroma. (A) Characteristic lymphoid stroma-associated markers gp38 (red), ER-TR7 (green), and LYVE-1 (cyan) in nondraining axillary lymph nodes and peripheral stromal zones of CCL21^{low}, control, and CCL21^{high} tumors by confocal microscopy at day 21 p.i. (B) Secretion of CCL21 (green) by gp38⁺ (red) ER-TR7⁺ (cyan) stromal cells in lymph nodes and tumors. (C) IDO (green) production in the gp38⁺ (red) tumor stroma. (D) Expression of the complement regulating protein Crry (red) within stromal (ER-TR7, green) compartments of lymph nodes and tumors. (E) Blood (CD31⁺, red) and lymphatic (LYVE-1⁺, green) vessels in lymph nodes and tumors. In all images, nuclei are counterstained with 4',6'-diamidino-2-phenylindole (DAPI), and dotted lines denote tumor (T)–dermis (D) border. LN, lymph node; C, capsule; P, paracortex. Scale bars, 50 μ m. (F) Numbers of CD3 ϵ ⁺CD4⁺ROR γ t⁺ LTi cells detected in tumors from wild-type ($n \geq 3$) or CCR7^{−/−} mice (control tumors only, $n = 4$). Bars show medians \pm SE. (G) Volumes of control tumors day 9 p.i. from wild-type C57/BL6 ($n = 22$), wild-type 129/P2 ($n = 5$), and LTi-deficient *Rorc*(γ t)^{GFP/GFP} mice ($n = 4$). Data represent means \pm SEM.

cells detected in tumors from wild-type ($n \geq 3$) or CCR7^{−/−} mice (control tumors only, $n = 4$). Bars show medians \pm SE. (G) Volumes of control tumors day 9 p.i. from wild-type C57/BL6 ($n = 22$), wild-type 129/P2 ($n = 5$), and LTi-deficient *Rorc*(γ t)^{GFP/GFP} mice ($n = 4$). Data represent means \pm SEM.

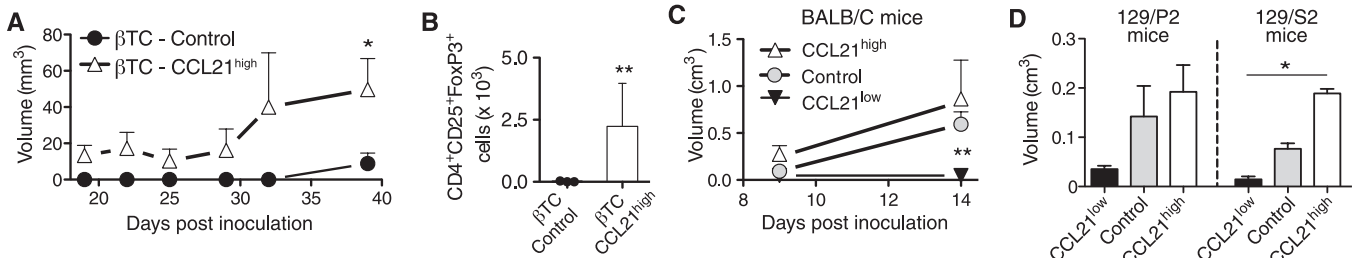


Fig. 4. CCL21 promotes survival of orthotopic and nonsyngeneic tumor allografts. **(A)** Growth rates for orthotopically implanted control-transfected and CCL21-overexpressing β tumor cells (β TC-control and β TC-CCL21^{high}; $n = 7$). Data represent means \pm SEM. **(B)** Intratumoral T_{reg} cells within control β tumors. Data represent means \pm SEM. **(C)** Volumes of CCL21^{low}, control, and CCL21^{high}

tumors in nonsyngeneic BALB/C recipients at days 9 and 14 p.i. ($n = 4$). Data represent means \pm SEM. **(D)** Volumes of CCL21^{low}, control, and CCL21^{high} tumors in nonsyngeneic 129/P2 and 129/S2 recipients 9 days p.i. ($n \geq 2$). Bars show medians \pm SE. * $P < 0.05$, ** $P < 0.01$ compared with relevant controls.

both in wild-type mice (Fig. 3F and fig. S6A) and *Rorc*(γ)^{+/GFP} mice, which generate GFP-expressing LTi cells (23) (fig. S6B). In contrast, CCL21-enhanced tumor growth was absent in LTi-deficient *Rorc*(γ)^{GFP/GFP} mice (Fig. 3G). Therefore, tumor expression of CCL21 was correlated with LTi-cell recruitment, although it is not clear whether LTi-cell recruitment was required for the CCL21-enhanced tumor growth and host immune tolerance observed.

This host tolerogenic response to CCL21-secreting tumors could also be demonstrated with another murine tumor cell line, islet beta tumor cells (BTCs) (27), which, when transduced to stably overexpress CCL21, grew significantly larger and contained more T_{reg} cells than control-transduced counterparts in syngeneic C57BL/6 mice (Fig. 4, A and B). In addition, CCL21 overexpression could even rescue nonsyngeneic allografts, including B16-F10 melanomas implanted into BALB/C, 129/P2 and S2 mice (Fig. 4, C and D). In these cases, we again found that, although CCL21^{low} tumors grew poorly, control and CCL21^{high} tumors grew robustly.

Taken together, these data suggest that CCL21 secretion by tumors led to a tolerogenic tumor microenvironment with stromal features resembling those of the LN paracortex. Consistent with recent findings that LN stroma itself plays an important role in promoting tolerance to self-antigens (9), we hypothesize that tumor CCL21-driven mimicry of the LN stroma helps promote a tolerogenic switch in the host immune response.

Several functions of CCL21 could help to drive the regulatory shift in the T cell populations that we saw in CCL21-expressing tumors. CCL21 can recruit naïve T cells to peripheral sites (15, 28) and can promote their differentiation into T_{reg} cells while inducing effector T cell senescence (7). The coincident development of the specialized stroma in CCL21-expressing tumors might enhance both T cell trafficking into the tumor and their interactions with APCs or even the stromal cells themselves (9). Such interactions within the regulatory cytokine environment of the tumor may promote T_{reg} cell activation and a further shift in the cytokine microenvironment toward one that is less immunogenic.

These changes are consistent with earlier reports that CCR7 signaling is required for the maintenance of peripheral self-tolerance (4, 7, 8), and with reports demonstrating promotion of deletional tolerance by the lymphoid stroma (9, 10). In contrast, CCL21 has also been associated with autoimmunity (7, 15), as it drives the formation of lymphocytic infiltrates and tertiary lymphoid structures. Such structures are characterized by B cell follicle formation (14); however, we did not detect B cell clusters in the tumors examined here (fig. S4F). These conflicting reports (7) emphasize that the timing and context in which CCR7⁺ leukocytes are recruited to the tumor, and the tumor cytokine environment, can modulate the outcome. Furthermore, it has been suggested that CCR7 ligands in the tumor might inhibit allograft rejection by entrapment of the APCs inside the tumor, which prevents them from migrating out to mount an immune response (19), and which our data do not support (fig. S2). Instead, we found CCL21-expressing tumors could prevent rejection of a nonsyngeneic allograft. We propose that, whereas CCL21-secreting tumors attract MDSCs, T_{reg} cells, and naïve T cells to the tumor microenvironment, it also induces lymphoid-like stroma that can equip the developing tumor with a substrate to promote the induction of T_{reg} cells and further guide naïve T cell interactions with APCs, all under a regulatory cytokine milieu (fig. S7). These findings hold therapeutic significance, particularly to tumor vaccine strategies and to emerging antitumor immunotherapies utilizing chemokines (including CCL21 and CCL19) (7) to functionally bias the recruited immune cell infiltrates within the tumor.

References and Notes

1. A. Ben-Baruch, *Cancer Metastasis Rev.* **25**, 357 (2006).
2. A. Mantovani, P. Allavena, A. Sica, F. Balkwill, *Nature* **454**, 436 (2008).
3. W. Zou, *Nat. Rev. Cancer* **5**, 263 (2005).
4. M. A. Schneider, J. G. Meingassner, M. Lipp, H. D. Moore, A. Rot, *J. Exp. Med.* **204**, 735 (2007).
5. A. Menning et al., *Eur. J. Immunol.* **37**, 1575 (2007).
6. J. C. Ochando et al., *J. Immunol.* **174**, 6993 (2005).
7. R. Förster, A. C. Davalos-Misslitz, A. Rot, *Nat. Rev. Immunol.* **8**, 362 (2008).

8. A. C. Davalos-Misslitz et al., *Eur. J. Immunol.* **37**, 613 (2007).
9. J. W. Lee et al., *Nat. Immunol.* **8**, 181 (2007).
10. S. N. Mueller, R. N. Germain, *Nat. Rev. Immunol.* **9**, 618 (2009).
11. A. Link et al., *Nat. Immunol.* **8**, 1255 (2007).
12. J. D. Shields et al., *Cancer Cell* **11**, 526 (2007).
13. T. D. Randall, D. M. Carragher, J. Rangel-Moreno, *Annu. Rev. Immunol.* **26**, 627 (2008).
14. D. L. Drayton, S. Liao, R. H. Mounzer, N. H. Ruddle, *Nat. Immunol.* **7**, 344 (2006).
15. W. Weninger et al., *J. Immunol.* **170**, 4638 (2003).
16. G. Eberl, D. R. Littman, *Immunol. Rev.* **195**, 81 (2003).
17. L. Peduto et al., *J. Immunol.* **182**, 5789 (2009).
18. S. Krautwald et al., *Immunology* **112**, 301 (2004).
19. E. Ziegler et al., *J. Am. Soc. Nephrol.* **17**, 2521 (2006).
20. Materials and methods are available as supporting material on Science Online.
21. B. Brierie, H. L. Moses, *Nat. Rev. Cancer* **6**, 506 (2006).
22. A. Sica et al., *Semin. Cancer Biol.* **18**, 349 (2008).
23. G. Eberl et al., *Nat. Immunol.* **5**, 64 (2004).
24. A. Kawase et al., *Int. J. Cancer* **123**, 1053 (2008).
25. G. C. Prendergast, *Oncogene* **27**, 3889 (2008).
26. J. C. Varela et al., *Cancer Res.* **68**, 6734 (2008).
27. J. A. Joyce et al., *Cancer Cell* **4**, 393 (2003).
28. M. R. Britschgi, A. Link, T. K. Lissandrini, S. A. Luther, *J. Immunol.* **181**, 7681 (2008).
29. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
30. We thank G. Eberl for *Rorc*(γ)^{+/GFP} and *Rorc*(γ)^{GFP/GFP} mice, S. Luther for CCR7^{-/-} mice, D. Hanahan for β tumor cells, and D. Trono for lentiviral vector plasmids; S. Pradervand, A. Paillisson, D. Foretay, M. Pasquier, V. Borel, B. Dixon, and A. Jimenez for technical assistance; and J. Hubbell and G. Eberl for critical reading of the manuscript. Funding was provided by the Swiss Cancer League, the Swiss National Science Foundation, the European Research Council, and the U.S. Department of Defense Breast Cancer Research Program to M.A.S. The authors (except J.M.R.) have filed a patent application on the use of CCL21 in immune modulation.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1185837/DC1
Materials and Methods
Figs. S1 to S7

11 December 2009; accepted 16 March 2010
Published online 25 March 2010;
10.1126/science.1185837
Include this information when citing this paper.

Altered Histone Acetylation Is Associated with Age-Dependent Memory Impairment in Mice

Shahaf Peleg,^{1*} Farahnaz Sananbenesi,^{1*} Athanasios Zovoilis,^{1*} Susanne Burkhardt,¹ Sanaz Bahari-Javan,¹ Roberto Carlos Agis-Balboa,¹ Perla Cota,¹ Jessica Lee Wittnam,^{1†} Andreas Gogol-Doering,² Lennart Opitz,³ Gabriella Salinas-Riester,³ Markus Dettenhofer,⁴ Hui Kang,² Laurent Farinelli,⁵ Wei Chen,² André Fischer^{1‡}

As the human life span increases, the number of people suffering from cognitive decline is rising dramatically. The mechanisms underlying age-associated memory impairment are, however, not understood. Here we show that memory disturbances in the aging brain of the mouse are associated with altered hippocampal chromatin plasticity. During learning, aged mice display a specific deregulation of histone H4 lysine 12 (H4K12) acetylation and fail to initiate a hippocampal gene expression program associated with memory consolidation. Restoration of physiological H4K12 acetylation reinstates the expression of learning-induced genes and leads to the recovery of cognitive abilities. Our data suggest that deregulated H4K12 acetylation may represent an early biomarker of an impaired genome-environment interaction in the aging mouse brain.

A number of studies indicate that aging correlates with brain region-specific changes of gene expression (1–4). It is, however, not well understood how aging affects gene expression and if those changes are causally linked to memory impairment. Remodeling of chromatin via histone acetylation, a key mechanism to control gene expression (5), has recently been implicated with the formation of long-term memories (6–10). Therefore, we hypothesized that altered histone acetylation might contribute to age-associated changes in gene expression and cognitive decline.

To detect an age at which cognitive impairment is first manifested, we subjected 3-, 8-, and 16-month-old C57BL/6 mice, which have a mean life span of 26 to 28 months (11, 12), to contextual fear conditioning (10), a commonly used test for hippocampus-dependent associative learning. Notably, the hippocampal formation is intimately involved in cognitive function in rodents and humans and is among the first to be affected during dementia (13). Whereas all groups were successfully able to learn this task, 16-month-old mice showed significantly less freezing behavior during the memory test, which indicated impaired associative learning (fig. S1A). Additional groups of mice were trained

in the Morris water-maze protocol (10), a well-established test for hippocampus-dependent spatial memory. All groups improved in their ability to find the hidden platform throughout the training trials, but the escape latency was significantly impaired in 16-month-old mice when compared with the 3- or 8-month-old groups (fig. S1B). Consistently, 16-month-old mice spent less time in

the target quadrant during a subsequent probe test (fig. S1C). Groups did not differ in finding a visible platform (fig. S1B), explorative behavior, or the response to a foot shock (fig. S1, D and E). Moreover, hippocampal levels of various markers for neuronal plasticity and integrity such as microtubule-associated protein 2, synaptophysin (Svp), postsynaptic density-95, synaptophysin (Svp), and glutamate receptor 1 were similar among 3- and 16-month-old mice (fig. S1, F and G). These findings suggest that the impairments in hippocampus-dependent memory formation displayed by 16-month-old mice cannot be explained by major structural changes, altered exploratory behavior, or impaired response to foot shock.

To test whether memory impairment correlates with altered chromatin plasticity, we first investigated whether hippocampal histone acetylation differs between 3- and 16-month-old naïve mice. Quantitative immunoblot analysis did not reveal significant changes of histone H3 acetylation on lysine residues (K) 9 and 14 or H4 acetylation on K5, 8, 12, or 16, which showed that the basal hippocampal histone acetylation profile is similar among 3- and 16-month-old naïve mice (fig. S2). In line with these data, the levels and activity of histone acetyltransferases (HATs) and histone deacetylases (HDACs) were similar in 3- and 16-month-old mice (fig. S3).

Recent data suggest that histone acetylation might play an important role in orchestrating the

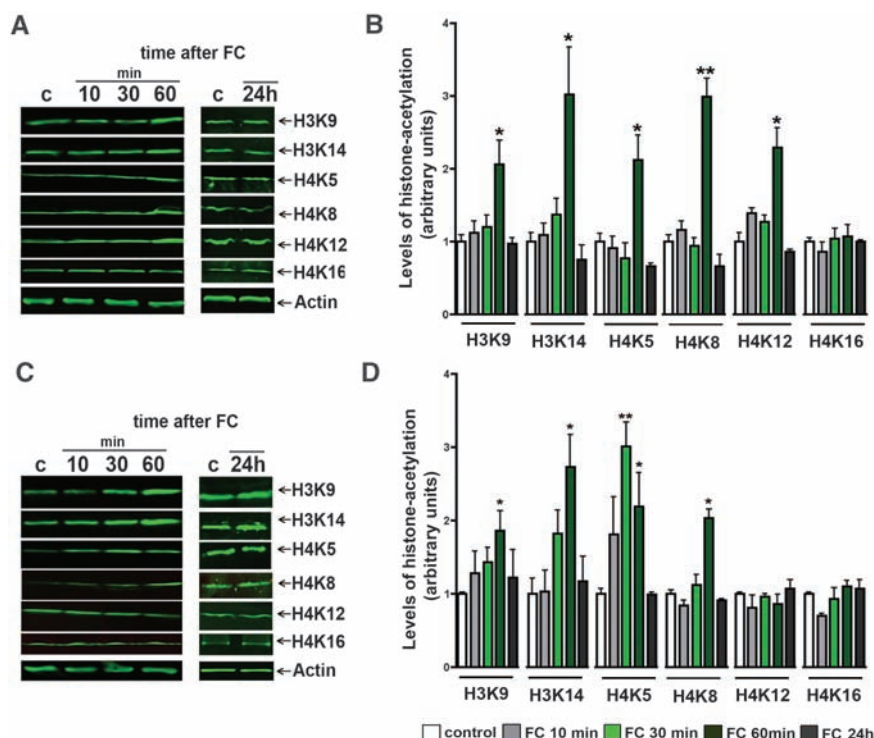


Fig. 1. Impaired learning and memory in 16-month-old mice correlates with deregulated H4K12 acetylation. (A) Representative immunoblot showing histone acetylation in 3-month-old mice in response to fear conditioning (FC). Control mice (c) were treated identically but did not receive the foot shock. (B) Quantification of (A). (C) Analysis similar to that described in (A) was performed in 16-month-old mice. (D) Quantification of (C). (** $P < 0.01$, * $P < 0.05$ versus control). $n =$ four or five mice per group. Error bars indicate SEM.

¹Laboratory for Aging and Cognitive Diseases, European Neuroscience Institute, Grisebach Str. 5, D-37077 Goettingen, Germany. ²Max Delbrueck Center for Molecular Medicine, Institute for Medical Systems Biology, Robert-Rössle-Strasse 10, D-13125 Berlin-Buch, Germany. ³DNA Microarray Facility, Georg August University, Humboldtallee 23, D-37073 Goettingen, Germany. ⁴Harvard Medical School, Genetics Department, 77 Ave Louis Pasteur, Boston, MA 02115, USA. ⁵Fasteris SA, CH-1228 Plan-les-Ouates, Switzerland.

*These authors contributed equally to this work.

†Present address: Department of Psychiatry, Division of Molecular Psychiatry, University Goettingen, von Siebold Str. 7, D-37075 Goettingen, Germany.

‡To whom correspondence should be addressed. E-mail: a.fischer@eni-g.de

gene expression program initiated by memory consolidation (8, 14). Therefore, we used quantitative immunoblotting to analyze hippocampal histone acetylation in 3- and 16-month-old mice at 10, 30, and 60 min and 24 hours after exposure to the fear conditioning training (fig. S4A). When compared with the age-matched control group, 3-month-old mice displayed a transient increase of H3K9 and H3K14 and H4K5, H4K8, and H4K12 acetylation 60 min after fear conditioning (Fig. 1, A and B). A similar transient increase of H3K9, H3K14, H4K5, and H4K8 acetylation was observed in 16-month-old mice. Whereas H4K5 acetylation was up-regulated 60 min after fear conditioning in both age groups, a significant increase of H4K5 acetylation was already detectable 30 min after fear conditioning in 16-month-old mice. These 16-month-old mice failed to up-regulate H4K12 acetylation (Fig. 1, C and D). These data were confirmed by immunohistochemical analysis (fig. S4). Moreover, the levels of total H4 did not change between groups (fig. S4), which showed that memory impairment

correlates with a deficit in learning-induced H4K12 acetylation in 16-month-old mice.

To analyze if deregulated H4K12 acetylation impacts learning-induced gene expression, we performed a high-density oligonucleotide microarray to compare the entire hippocampal gene expression profile of 3- and 16-month-old mice during memory consolidation (fig. S5A). To this end, 3- and 16-month-old mice were subjected to fear conditioning. Explorative behavior during the training and the response to the foot shock were similar among groups (fig. S5B). Animals that were not subjected to fear conditioning but otherwise were treated identically served as controls. Notably, the gene expression profile was nearly identical among 3- and 16-month-old control mice (Fig. 2A and fig. S6). This is consistent with our finding that histone acetylation, HAT, and HDAC activities are similar among those groups (figs. S2 and S3). In 3-month-old mice, 2229 genes (1980 up-regulated versus 449 down-regulated) were differentially expressed 1 hour after fear conditioning as compared with the age-

matched control group (Fig. 2A and table S1). However, the hippocampal transcriptome of 16-month-old mice remained almost unchanged in response to fear conditioning. When compared with the age-matched control group, only six genes were differentially expressed among groups (Fig. 2A and fig. S7).

Further analysis revealed that, in 3-month-old mice, 1539 of the differentially expressed genes were specifically linked to associative learning, hereafter called “learning-regulated genes” (fig. S8 and tables S2 and S3). The learning-regulated genes were associated with biological processes such as transcription, protein modification, or intracellular signaling (fig. S5C and table S4). Additional data mining revealed that 3-month-old mice regulate key signaling pathways implicated with memory formation and synaptic remodeling in response to fear conditioning (fig. S9 and table S5). Using quantitative polymerase chain reaction (qPCR) analysis, we confirmed the differential expression of genes selected to represent signaling pathways identified by data

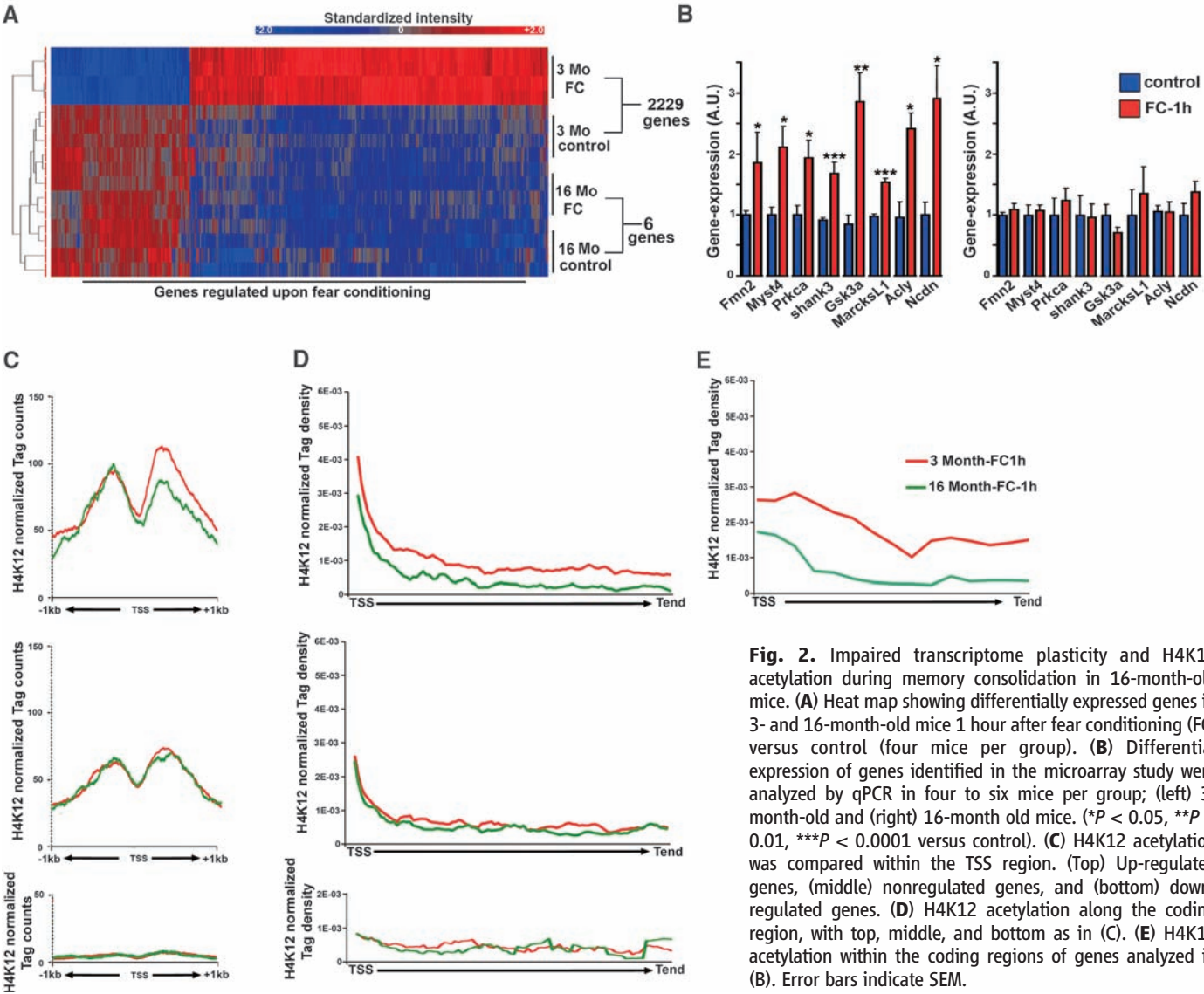


Fig. 2. Impaired transcriptome plasticity and H4K12 acetylation during memory consolidation in 16-month-old mice. (A) Heat map showing differentially expressed genes in 3- and 16-month-old mice 1 hour after fear conditioning (FC) versus control (four mice per group). (B) Differential expression of genes identified in the microarray study were analyzed by qPCR in four to six mice per group; (left) 3-month-old and (right) 16-month old mice. (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.0001$ versus control). (C) H4K12 acetylation was compared within the TSS region. (Top) Up-regulated genes, (middle) nonregulated genes, and (bottom) down-regulated genes. (D) H4K12 acetylation along the coding region, with top, middle, and bottom as in (C). (E) H4K12 acetylation within the coding regions of genes analyzed in (B). Error bars indicate SEM.

mining (Fig. 2B and table S6). Our data reveal that 16-month-old mice show severe impairment in regulating gene expression on exposure to relevant environmental stimuli that initiate learning behavior. As a result, key signaling pathways initiated in 3-month-old mice during associative learning are not properly regulated in 16-month-old mice.

Next, we investigated how altered H4K12 acetylation contributes to the lack of learning-induced gene expression in 16-month-old mice. We decided to take a genome-wide approach to analyze H4K12 acetylation of learning-regulated genes in 3- and 16-month-old mice after fear-conditioning. To this end, we used the recently developed ChIP-seq technology (15) that depends on the cross-linking of proteins to specific DNA elements, followed by immunoprecipitation of the protein-DNA complex and high-throughput sequencing of the recovered DNA (16, 17). Hippocampal tissue isolated from 3- and 16-month-old mice 1 hour after fear conditioning was subjected to H4K12 ChIP-seq (fig. S10). To generate a genome-wide map of H4K12 acetylation in the hippocampus of young and old mice during memory formation, the resulting reads were mapped to a reference mouse genome (fig. S10).

We first examined H4K12 acetylation in the region spanning the transcription start site (TSS),

which is essential for transcriptional initiation (16). To this end, we compared enrichment in H4K12 acetylation in 3- and 16-month-old mice at regions extending 1 kb upstream and 1 kb downstream of the TSS (hereafter, referred to as the TSS region) of (1) up-regulated genes, (2) randomly chosen genes that were not regulated upon fear conditioning, and (3) learning-regulated genes that were down-regulated (Fig. 2C, top, middle, and bottom, respectively). H4K12 acetylation of down- or nonregulated genes did not differ among groups (Fig. 2C). However, H4K12 acetylation of up-regulated genes was altered in 16-month-old mice (Fig. 2C, top). Impaired H4K12 acetylation was specifically observed in the genomic region 1 kb downstream of the TSS (Fig. 2C), which marks the beginning of the gene-coding region. H4K12 acetylation 1 kb upstream of the TSS, a region that marks the gene promoter, was similar among groups. Therefore, we analyzed the distribution of H4K12 acetylation within the coding regions of up-, non-, and down-regulated genes, as well as the genes individually analyzed in Fig. 2B (Fig. 2, D and E, and fig. S11). The levels of H4K12 acetylation in up-regulated genes were lower in 16-month-old mice than in the 3-month-old group. No difference was observed for non- and down-regulated genes (Fig. 2D).

ChIP-seq analysis of H3K9 acetylation did not reveal higher enrichment in 3-month-old compared with 16-month-old mice (fig. S12), which is in line with our findings that H4K12, but not H3K9, acetylation is deregulated in 16-month-old mice on fear conditioning (see also Fig. 1). A recent study showed that in blood cells H4K12 is mainly enriched within gene bodies and therefore associated with transcriptional elongation, whereas other sites, such as H3K9, peak in the TSS region (16). In line with this, we found that high levels of gene expression in the hippocampus also correlate with high levels of H4K12 acetylation along the coding regions of genes, whereas no such correlation was observed for H3K9 (fig. S13). These data suggest that the severe lack of learning-induced gene expression in 16-month-old mice is linked, at least in part, to deregulated H4K12 acetylation associated with impaired transcriptional elongation of up-regulated genes.

To analyze this in greater detail, we investigated the Formin 2 gene, *Fmn2*, as an example. Formin 2 is an actin nucleator highly expressed in the adult brain (18), and actin dynamics are of great importance for synaptic plasticity and memory formation (19, 20). Moreover, *Fmn2* was induced in 3-month-old, but not in 16-month-old, mice on fear conditioning (Fig. 2B), which correlated with impaired H4K12 acetylation

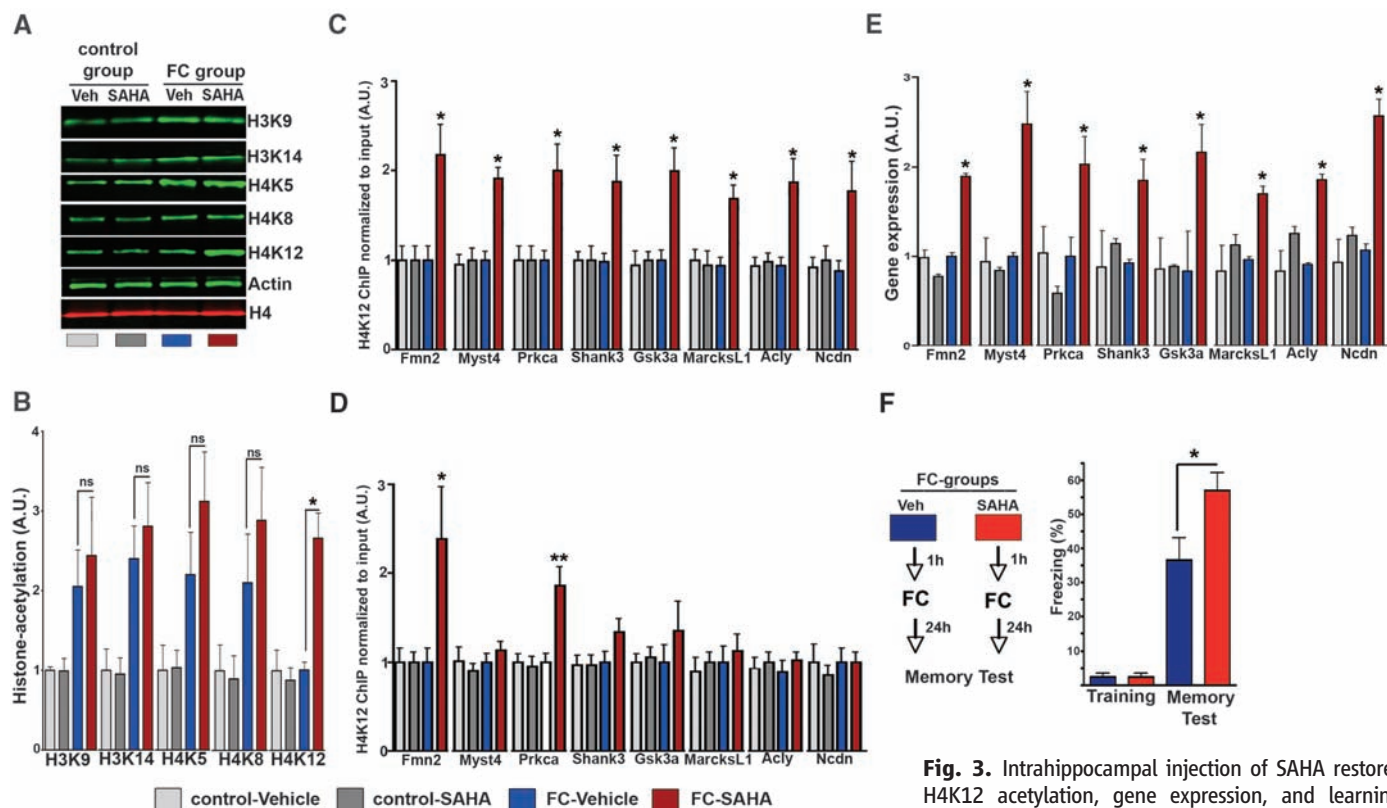


Fig. 3. Intrahippocampal injection of SAHA restores H4K12 acetylation, gene expression, and learning behavior in 16-month-old mice. (A) Representative immunoblots showing the levels of hippocampal histone acetylation in vehicle or SAHA-treated mice. (B) Quantification of (A). H4K12 acetylation was significantly increased in the SAHA-treated mice after fear conditioning ($P < 0.05$, four mice per group). A.U., arbitrary units. (C) H4K12 ChIP of coding region. SAHA-treatment increased learning-induced H4K12 acetylation in the coding regions of all investigated genes ($P < 0.05$, four mice per group). (D) H4K12 ChIP of promoter region. SAHA-treatment increased learning-induced H4K12 acetylation in the formin 2 and *Prkca* promoter ($P < 0.05$, four mice per group). (E) Gene expression. Increased learning-induced gene expression in 16-month-old mice treated with SAHA ($P < 0.05$, five or six mice per group). (F) Associative learning was enhanced in SAHA-treated mice ($P < 0.05$, nine mice per group). Error bars indicate SEM.

immunoblot showing the levels of hippocampal histone acetylation in vehicle or SAHA-treated mice. (B) Quantification of (A). H4K12 acetylation was significantly increased in the SAHA-treated mice after fear conditioning ($P < 0.05$, four mice per group). A.U., arbitrary units. (C) H4K12 ChIP of coding region. SAHA-treatment increased learning-induced H4K12 acetylation in the coding regions of all investigated genes ($P < 0.05$, four mice per group). (D) H4K12 ChIP of promoter region. SAHA-treatment increased learning-induced H4K12 acetylation in the formin 2 and *Prkca* promoter ($P < 0.05$, four mice per group). (E) Gene expression. Increased learning-induced gene expression in 16-month-old mice treated with SAHA ($P < 0.05$, five or six mice per group). (F) Associative learning was enhanced in SAHA-treated mice ($P < 0.05$, nine mice per group). Error bars indicate SEM.

tion throughout the coding region (Fig. 2E and fig. S11). We confirmed this finding using ChIP followed by qPCR analysis to show that H4K12 acetylation of *Fmn2* increases 1 hour after fear conditioning in 3-month-old, but not in 16-month-old, mice (fig. S14, A and B). The observed changes in H4K12 acetylation and mRNA expression translated into differential protein production. Formin 2 protein levels transiently increased in response to fear-conditioning in 3-month-old, but not in 16-month-old, mice (fig. S14C).

Next, we used mice lacking *Fmn2* (21) to investigate its role in memory formation. These mice are viable and show normal brain anatomy (fig. S14D). Associative learning was similar among 3-month-old *Fmn2*^{−/−} mice and wild-type littermates. However, 8-month-old *Fmn2*^{−/−} mice showed impaired associative learning ability when compared with an age-matched control group (fig. S14E). Pain sensation, explorative behavior, and basal anxiety were similar among groups (fig. S14, F to H). These data suggest that reduced formin 2 levels contribute to age-related memory impairment.

Nevertheless, the age-associated memory impairment in *Fmn2*^{−/−} mice cannot be directly compared with memory impairment observed in wild-type mice. In 16-month-old wild-type mice expression of *Fmn2*, as well as those of the 1538 other learning-regulated genes, is normal under basal conditions. However, these mice fail to increase the levels of learning-induced genes during associative learning. Therefore, we propose that rather than targeting the expression levels of a single candidate gene, elevation of H4K12 acetylation during aging might be a more efficient approach to restore learning abilities.

To test this hypothesis, we implanted microcannulae into the hippocampi of 16-month-old mice. One group of mice was injected with the potent HDAC inhibitor suberoylanilide hydroxamic acid (SAHA, 10 μg per hippocampus); the other group received vehicle solution. One hour after injection, mice were subjected to contextual fear conditioning and killed 1 hour later for molecular analysis. Mice that were injected with SAHA or vehicle and used for molecular analysis 2 hours later without being subjected to fear conditioning served as additional control groups (fig. S15A). Notably, SAHA-treated mice showed significantly increased hippocampal H4K12 acetylation after fear conditioning (Fig. 3, A and B). In line with this finding, we observed that, after fear conditioning, SAHA-treated mice displayed elevated H4K12 acetylation in the coding regions of learning-regulated genes (Fig. 3C). The effect of SAHA on H4K12 acetylation in the promoter regions was less pronounced, which further supported a predominant role of H4K12 in transcriptional elongation (Fig. 3D). The expression of learning-regulated genes (table S5) was significantly higher in 16-month-old mice that were treated with SAHA before fear conditioning (Fig. 3E).

These data suggest that increasing H4K12 acetylation restores learning-induced gene ex-

pression. This work may show relevance therapeutically for the treatment of age-associated memory impairment via the restoration of physiological expression levels for genes contributing to memory consolidation. Indeed, 16-month-old mice treated with SAHA showed facilitated associative learning when compared with the vehicle group (Fig. 3F). The response to the foot shock, explorative behavior, and tone-dependent fear conditioning that is hippocampus independent were not affected (fig. S15, B to D), which showed that the SAHA-mediated increase of H4K12 acetylation is sufficient to restore associative learning in 16-month-old mice. Similar data were obtained using the pan-HDAC inhibitor sodium butyrate (fig. S16), whereas administration of an HDAC inhibitor that failed to increase H4K12 acetylation in 16-month-old mice also failed to reinstate learning ability (fig. S17).

In summary, our data suggest that deregulated H4K12 acetylation is causally involved in age-associated memory impairment. Although we cannot exclude that other histone modifications also contribute to this effect, recent studies support a unique role of H4K12 in the orchestration of gene expression (16). As such, H4K12 acetylation seems to be of particular importance for transcriptional elongation that is characterized by high levels of histone modifications along gene bodies (16, 22). Consistently, we show that deregulation of H4K12 acetylation in 16-month-old mice is mainly found along gene bodies of up-regulated genes. This may also explain why a deficit in H4K12 acetylation could mediate the observed profound effect on learning-regulated gene expression. As such, even in the presence of proper transcriptional initiation, further up-regulation of learning-regulated genes above baseline levels would be impaired when transcriptional elongation is affected. The precise mechanisms that underlie the selective deregulation of H4K12 acetylation in 16-month-old mice remain to be elucidated. It is likely that during aging a combination of multiple factors contributes to deregulated histone acetylation. Although we could not detect significant differences between changes in hippocampal HAT or HDAC activity in 3-month-old versus 16-month-old mice (figs. S3 and S18), this does not exclude the possibility that specific HATs or HDACs contribute to deregulated gene expression and learning impairment in 16-month-old mice. In fact, the HATs *Myst4* and *Gcn5l2*, as well as HDAC2 and HDAC4, were differentially expressed in 3-month-old, but not in 16-month-old, mice during memory formation (fig. S18 and table S3). Recent data showed that HDAC2 regulates memory formation (23). Moreover, small changes in the metabolic state of the cell may favor the dysfunction of mechanisms that are engaged with H4K12 acetylation. For example, a recent study demonstrated that histone acetylation critically depends on citrate levels (24), which are reduced in the aging brain (25) (fig. S19).

In conclusion, we found that the administration of HDAC inhibitors that shift the balance of H4K12 acetylation is able to reinstate learning-induced gene expression and memory function in 16-month-old mice. Our data also suggest that H4K12 acetylation-dependent changes in gene expression may serve as an early biomarker for an impaired genome-environment interaction in the aging brain.

References and Notes

- X. Xu et al., *Genome Biol.* **8**, R234 (2007).
- T. Lu et al., *Nature* **429**, 883 (2004).
- N. C. Berchtold et al., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 15605 (2008).
- C. K. Lee, R. G. Klopp, R. Weindrich, T. A. Prolla, *Science* **285**, 1390 (1999).
- T. Jenuwein, C. D. Allis, *Science* **293**, 1074 (2001).
- J. M. Alarcón et al., *Neuron* **42**, 947 (2004).
- E. Korzus, M. G. Rosenfeld, M. Mayford, *Neuron* **42**, 961 (2004).
- J. M. Levenson et al., *J. Biol. Chem.* **279**, 40545 (2004).
- C. G. Vecsey et al., *J. Neurosci.* **27**, 6128 (2007).
- A. Fischer, F. Sananbenesi, X. Wang, M. Dobbin, L. H. Tsai, *Nature* **447**, 178 (2007).
- M. Jucker, D. K. Ingram, *Behav. Brain Res.* **85**, 1 (1997).
- Materials and methods are available as supporting material on Science Online.
- M. M. Mesulam, *Neuron* **24**, 521 (1999).
- W. B. Chwang, J. S. Arthur, A. Schumacher, J. D. Sweatt, *J. Neurosci.* **27**, 12732 (2007).
- E. R. Mardis, *Nat. Methods* **4**, 613 (2007).
- Z. Wang et al., *Nat. Genet.* **40**, 897 (2008).
- A. Visel et al., *Nature* **457**, 854 (2009).
- B. Leader, P. Leder, *Mech. Dev.* **93**, 221 (2000).
- A. Fischer, F. Sananbenesi, C. Schrick, J. Spiess, J. Radulovic, *J. Neurosci.* **24**, 1962 (2004).
- Y. Fukazawa et al., *Neuron* **38**, 447 (2003).
- B. Leader et al., *Nat. Cell Biol.* **4**, 921 (2002).
- D. C. Hargreaves, T. Horng, R. Medzhitov, *Cell* **138**, 129 (2009).
- J. S. Guan et al., *Nature* **459**, 55 (2009).
- K. E. Wellen et al., *Science* **324**, 1076 (2009).
- N. Jiang et al., *J. Proteome Res.* **7**, 3678 (2008).
- We thank J. Radulovic, S. Irniger, A. Kranz, and W. Fischle for reading the manuscript and helpful comments. Microarray and ChIP-seq data are accessible through GEO Series accession number GSE20270 (www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE20270) and Galaxy data-sharing platform (<http://main.g2.bx.psu.edu/u/fischerlab/h/sm1186088>; published history sm1186088). This work was partially supported by the following funds to A.F.: The EURY award of the European Science Foundation, the Hans and Ilse Breuer Foundation, the Schramm Foundation, and the European Research Area (ERA)-Net Neuron project Epithery. S.P. is supported by a Minerva fellowship, F.S. is supported by the Deutsche Forschungsgemeinschaft (German Research Foundation), P.C. is supported by a German Academic Exchange Service (DAAD) fellowship, and R.C.A.-B. is supported by a European Molecular Biology Organization (EMBO) long-term fellowship. M.D. is a Howard Hughes Medical Institute postdoctoral fellow, and L.F. is a founder and shareholder of FASTER SA. The European Neuroscience Institute is jointly funded by the University Medicine Goettingen and the Max Planck Society.

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/753/DC1

Materials and Methods

Figs. S1 to S19

Tables S1 to S7

References

17 December 2009; accepted 10 March 2010
10.1126/science.1186088

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/757>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/328/5979/757/DC1>

This article **cites 28 articles**, 10 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/328/5979/757#otherarticles>

This article appears in the following **subject collections**:

Cell Biology

http://www.sciencemag.org/cgi/collection/cell_biol

Molecular Biology

http://www.sciencemag.org/cgi/collection/molec_biol

Sequential Checkpoints Govern Substrate Selection During Cotranslational Protein Targeting

Xin Zhang, Rumana Rashid,* Kai Wang,† Shu-ou Shan‡

Proper protein localization is essential for all cells. However, the precise mechanism by which high fidelity is achieved is not well understood for any protein-targeting pathway. To address this fundamental question, we investigated the signal recognition particle (SRP) pathway in *Escherichia coli*, which delivers proteins to the bacterial inner membrane through recognition of signal sequences on cargo proteins. Fidelity was thought to arise from the inability of SRP to bind strongly to incorrect cargos. Using biophysical assays, we found that incorrect cargos were also rejected through a series of checkpoints during subsequent steps of targeting. Thus, high fidelity of substrate selection is achieved through the cumulative effect of multiple checkpoints; this principle may be generally applicable to other pathways involving selective signal recognition.

Cotranslational protein targeting is an essential and evolutionarily conserved pathway for delivering proteins to the eukaryotic endoplasmic reticulum, or the bacterial plasma membrane (1, 2). Targeting begins when the signal recognition particle (SRP) recognizes an N-terminal signal sequence on its cargo, a translating ribosome bearing a nascent polypeptide chain (RNC) (Fig. 1A, step 1) (1–6). Cargo-loading

facilitates efficient interaction between the guanine triphosphatase (GTPase) domains of both the SRP and SRP receptor (SR) and stabilizes the SRP-SR complex in an early conformational state (step 2) (7, 8). The interactions of SR with the target membrane and the protein-conducting channel are proposed to induce dynamic rearrangements in the SRP-SR complex (4, 6, 8), first to form a guanosine triphosphate (GTP)-dependent closed complex (step 3), and then to activate GTP hydrolysis in the complex (step 4). These rearrangements facilitate the unloading of cargo from SRP to the translocation machinery (steps 3 to 4) (4, 6, 8). In a productive targeting cycle, GTP is hydrolyzed after cargo unloading to drive the disassembly and recycling of SRP and SR (step 5) (9).

How SRP ensures faithful delivery of correct cargos remains poorly understood. The SRP

signal sequences are highly degenerate, and their difference from the signal sequences of non-SRP substrates is minor (10). Thus, SRP must have evolved a strategy to remain highly specific to its substrates despite the “noise” in its recognition signal. “Incorrect” cargos were thought to be rejected because they bind weakly to the SRP (11) (Fig. 1A, arrow a). To test this hypothesis, we systematically varied the signal sequence based on alkaline phosphatase (phoA) (10, 12–14). We replaced the hydrophobic core of the phoA signal sequence (Fig. 1B) with a combination of leucine and alanine, and varied the Leu/Ala ratio to generate signal sequences with different hydrophobicity (12, 13, 15). Another incorrect cargo is the *Escherichia coli* autotransporter EspP. Although EspP contains a signal sequence with hydrophobicity comparable to that of phoA-3A7L, it is not an SRP substrate because it has an unusual N-terminal extension (Fig. 1B) (16). Firefly luciferase, a cytosolic protein without signal sequences, was used as a negative control (Fig. 1B) (12). For all the experiments, homogeneous stalled RNCs were purified and used as cargos (8, 17).

We first tested the binding affinities of SRP for different cargos (Fig. 1A, step 1). RNC binding to SRP was detected as an increase in the fluorescence anisotropy of fluorescein-labeled SRP (C421). Cargos with the most hydrophobic signal sequences bound to SRP tightly (RNC_{1A9L} and RNC_{2A8L}), with equilibrium dissociation constants (K_d) of ~1 nM or less (Fig. 1C and fig. S1). The next strongest cargo, RNC_{3A7L}, also exhibited strong, albeit attenuated binding to SRP, with K_d ~10 nM (Fig. 1C), and cargos weaker than RNC_{phoA} bound SRP with K_d values higher by one order of magnitude. Nevertheless, the affinity of incorrect

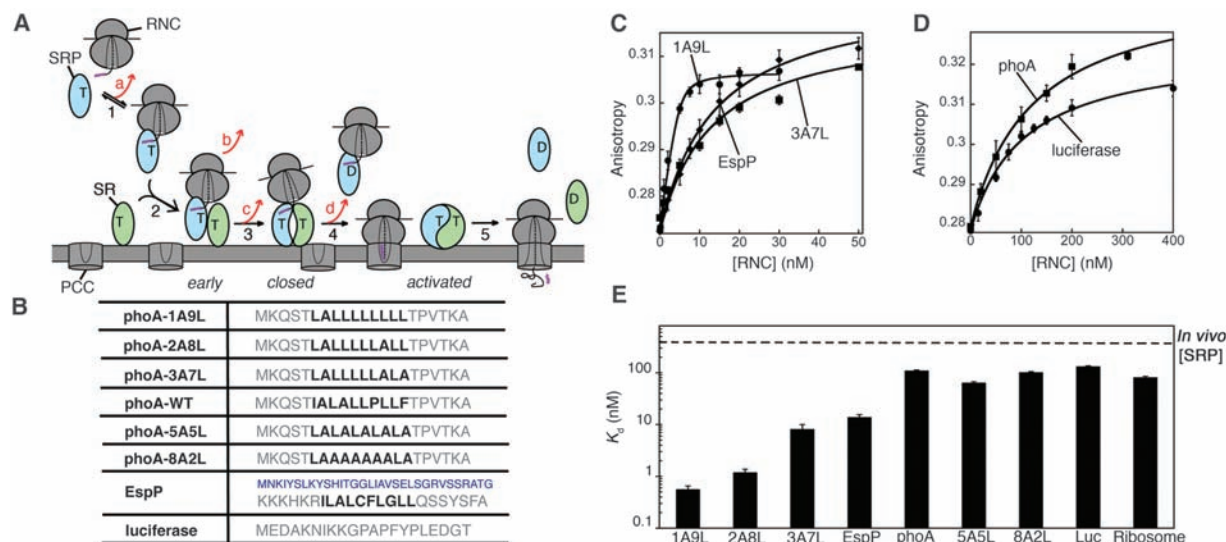


Fig. 1. Potential fidelity checkpoints in the SRP pathway. (A) Model for potential checkpoints during cotranslational protein targeting. A cargo (RNC) with a signal sequence (magenta) enters the pathway upon binding SRP and is either retained (black arrows) or rejected (red arrows) at each step (numbered 1 to 5). T and D denote GTP and GDP, respectively. (B) Signal sequence variants used in this study. Bold highlights the hydrophobic core. Blue highlights the N-terminal signal sequence extension of

EspP. (C and D) Equilibrium titrations of SRP-RNC binding. (C) RNC_{1A9L} (●), RNC_{3A7L} (■), and RNC_{EspP} (◆). (D) RNC_{phoA} (■) and RNC_{luciferase} (●). K_d values were determined by nonlinear fits of data and are summarized in (E) and table S1. Error bars are SDs from three independent experiments. (E) Summary of the binding affinities of SRP for different cargos. The dashed line denotes the cellular SRP concentration of 400 ± 58 nM. Error bars are SDs from three independent experiments.

cargos or the empty ribosome for SRP was still substantial, with K_d values of 80 to 100 nM (Fig. 1, D and E, and fig. S1) (18). At the cellular SRP concentration of ~400 nM (Fig. 1E) (19), a substantial amount of incorrect cargos could bind to SRP (20). Although EspP is not an SRP substrate, RNC_{EspP} unexpectedly bound SRP as tightly as RNC_{3A7L} (Fig. 1C). Thus, the differences in cargo binding affinity do not provide sufficient discrimination against incorrect cargos, and additional factors in the bacterial cytosol did not increase the specificity of SRP-cargo binding [fig. S2 and supporting online material (SOM) text] (21). Hence, subsequent steps in the targeting pathway, including formation of the SRP-SR complex and GTP hydrolysis, may provide additional checkpoints to reject the incorrect cargos (Fig. 1A, red arrows b to d and SOM text) (22).

We next tested whether the early SRP-SR complex is stabilized more strongly by the correct than the incorrect cargo (Fig. 1A, arrow b). We assembled cargo-SRP-SR early complexes in the absence of nucleotides; this blocks the rearrangement of the GTPase complex to subsequent conformations (7, 8). The equilibrium stabilities of the early complexes were measured using fluorescence resonance energy transfer (FRET) between donor- and acceptor-labeled SRP and SR (7). The early complex was significantly stabilized by RNC_{1A9L} and RNC_{2A8L}, with a K_d of ~80 nM (Fig. 2A), and this stability

was lowered by a factor of up to 50 for the weaker cargos (Fig. 2, B and C, and fig. S3). With incorrect cargos such as RNC_{EspP} and RNC_{luciferase}, the FRET efficiency also plateaued at lower values, ~0.3 to 0.4 (Fig. 2, B and D, and fig. S3), compared with FRET efficiency of ~0.66 with the correct cargos (Fig. 2, A and D). Thus the SRP and SR are likely to be positioned differently in the early targeting complexes formed by the incorrect cargos.

A mispositioned early complex would lead to a slower rearrangement to form the closed complex (Fig. 1A, step 3). To test this hypothesis, we preformed the early targeting complex and directly measured its rearrangement using acrylodan-labeled SRP (C235), which specifically monitors the closed complex. With RNC_{1A9L}, this rearrangement occurred at 0.3 s⁻¹ (Fig. 2E). RNC_{3A7L} and RNC_{phoA} mediated this rearrangement with rate constants 40% slower than that of RNC_{1A9L} (Fig. 2G and fig. S4). Notably, the rate constants at which RNC_{EspP} and cargos weaker than RNC_{5A5L} mediated this rearrangement were slower than that of RNC_{1A9L} by a factor of 5 to 10 (Fig. 2, F and G, and fig. S4). Thus, incorrect cargos do not induce the formation of a stable and productive early complex and are more likely to exit the pathway prematurely (Fig. 1A, arrow b).

The more favorable preequilibrium to form the early intermediate combined with the faster

early-to-closed rearrangement would allow the correct cargos to mediate faster GTP-dependent assembly of a stable closed complex (Fig. 1A, steps 2 and 3). We characterized this cumulative effect using both FRET (Fig. 3, A to C, and fig. S5, F and G) and acrylodan-labeled SRP (C235) (fig. S5). Both probes showed that the correct cargos mediated rapid assembly of the closed complex (Fig. 3A and fig. S5A), and this rate decreased significantly with weaker signal sequences (Fig. 3, B and C, and fig. S5). Overall, there is a ~10³-fold kinetic discrimination between the strongest and weakest cargos in stable SRP-SR complex assembly, which delivers the cargo to the membrane (Fig. 3C and fig. S5E).

At the membrane, GTP hydrolysis could compete with the unloading of cargo. If GTP hydrolyzed too quickly in the SRP-SR complex, this would abort the targeting reaction before the cargo is productively unloaded (8, 23). To test whether the correct cargos prevent premature GTP hydrolysis better than the incorrect cargos (Fig. 1A, step 4), we determined the GTPase rates from the cargo-SRP-SR complex. RNC_{1A9L} and RNC_{2A8L} reproducibly delayed GTP hydrolysis by a factor of 6 to 8 compared with the rate in the absence of cargo (Fig. 3D and fig. S6). RNC_{3A7L} also had an inhibitory effect on the GTPase reaction, reducing the GTPase rate by a factor of 3 to 4 (fig. S6). In contrast, incorrect cargos such as RNC_{EspP} did not substantially affect the GTPase

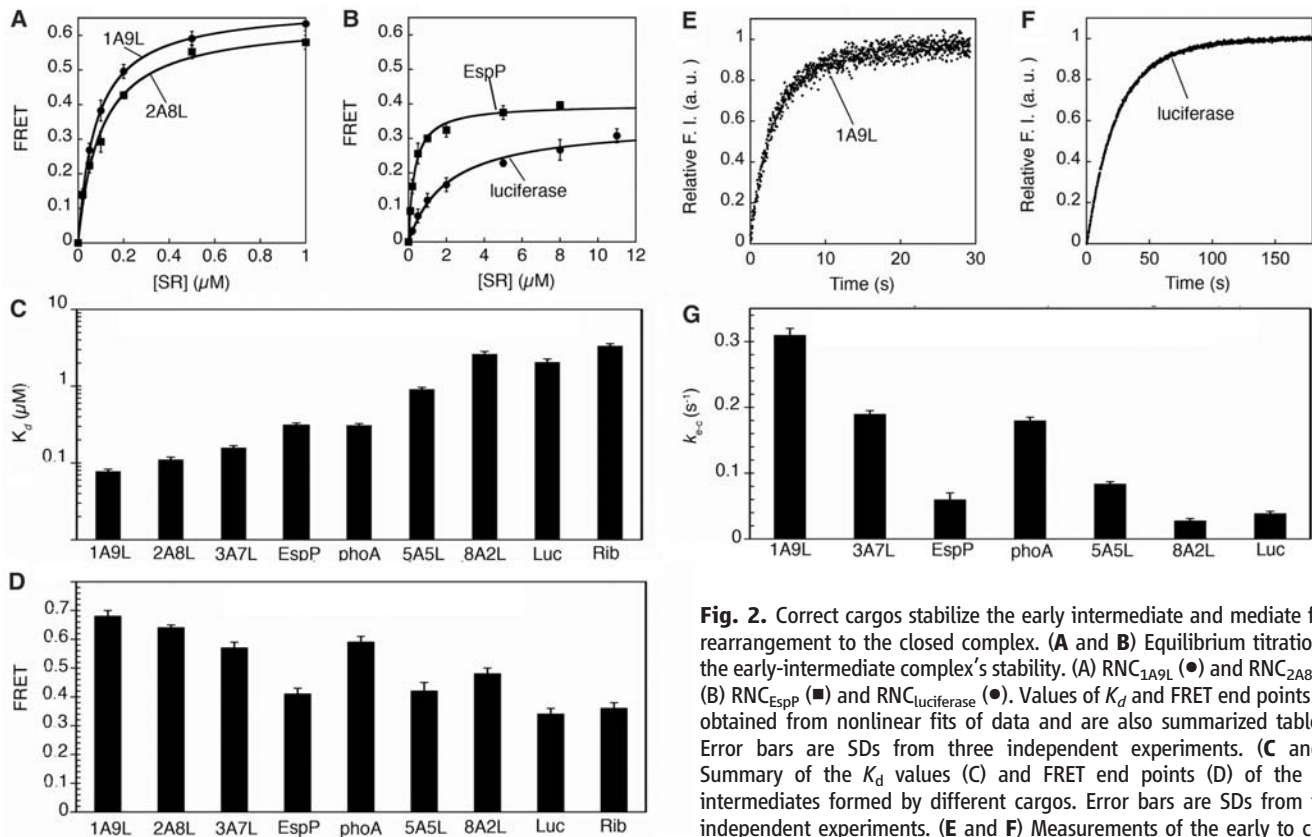


Fig. 2. Correct cargos stabilize the early intermediate and mediate faster rearrangement to the closed complex. (A and B) Equilibrium titrations of the early-intermediate complex's stability. (A) RNC_{1A9L} (●) and RNC_{2A8L} (■); (B) RNC_{EspP} (■) and RNC_{luciferase} (●). Values of K_d and FRET end points were obtained from nonlinear fits of data and are also summarized table S1. Error bars are SDs from three independent experiments. (C and D) Summary of the K_d values (C) and FRET end points (D) of the early intermediates formed by different cargos. Error bars are SDs from three independent experiments. (E and F) Measurements of the early to closed rearrangement. The rearrangement rate constants were obtained from exponential fits of data and are summarized in (G) and table S1. Error bars are SDs from three independent experiments. (G) Summary of the rate constants for the early to closed rearrangement with different cargos. Error bars are SDs from three independent experiments.

exponential fits of data and are summarized in (G) and table S1. Error bars are SDs from three independent experiments. (G) Summary of the rate constants for the early to closed rearrangement with different cargos. Error bars are SDs from three independent experiments.

rate (Fig. 3, E and F, and fig. S6). Thus, the fidelity of protein targeting can be further improved through kinetic proofreading mechanisms by using the energy of GTP hydrolysis (Fig. 1A, arrow d).

Our data suggested a model in which the incorrect cargos are rejected not only through binding affinity, but also through differences in the kinetics of SRP-SR complex assembly and GTP hydrolysis (Figs. 1A and 4A, top). On the basis of this model, we calculated the amount of substrates retained in the SRP pathway after each checkpoint (14). The cargo-binding step was not sufficient to discriminate against incorrect cargos, allowing over 75% of them to enter the SRP pathway (Fig. 4A, light gray). During cargo delivery through SRP-SR complex assembly, a large fraction of substrates weaker than *phoA* were rejected

(Fig. 4A, dark gray). Finally, kinetic competition between GTP hydrolysis and cargo unloading further improved the discrimination between correct and incorrect substrates (Fig. 4A, black). To validate the model, we determined the targeting efficiency of proteins with various signal sequences, using a well-established assay that tests the ability of *E. coli* SRP and SR to mediate the cotranslational targeting of preproteins to microsomal membranes (14, 24, 25). Substrates with signal sequences stronger than 3A7L were efficiently targeted and translocated (Fig. 4B and fig. S7). In contrast, *EspP* and substrates with signal sequences weaker than *phoA* showed severe defects in translocation, and almost no translocation was detected for *phoA*-8A2L (Fig. 4B). The experimentally determined protein-targeting effi-

ciencies agreed well with predictions based on the kinetic and thermodynamic measurements (Fig. 4C), which suggested that our model (Figs. 1A and 4A) faithfully represents how SRP handles its substrates.

Thus, fidelity during cotranslational protein targeting is achieved through the cumulative effect of multiple checkpoints, by using a combination of binding, induced fit, and kinetic proofreading mechanisms (SOM text). Although the incorrect cargos are not completely rejected during the initial binding step, they are discriminated repeatedly during subsequent steps possibly because they bind the SRP in a less productive mode (6). In addition, the translocation machinery provides another important checkpoint (26). Similar strategies of using multiple

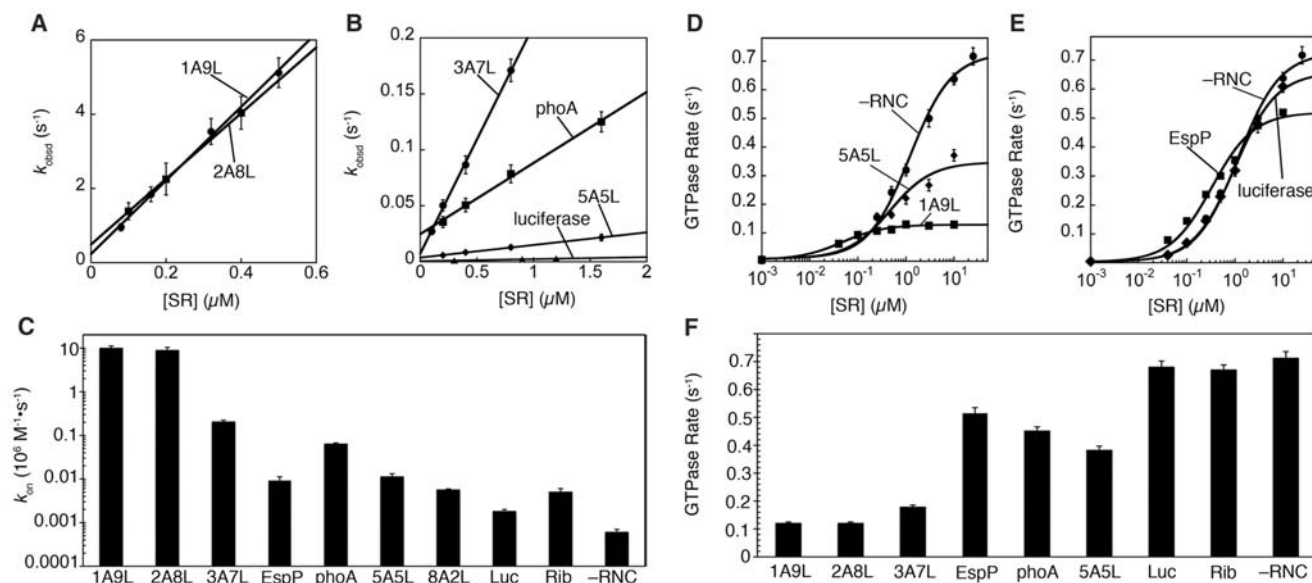
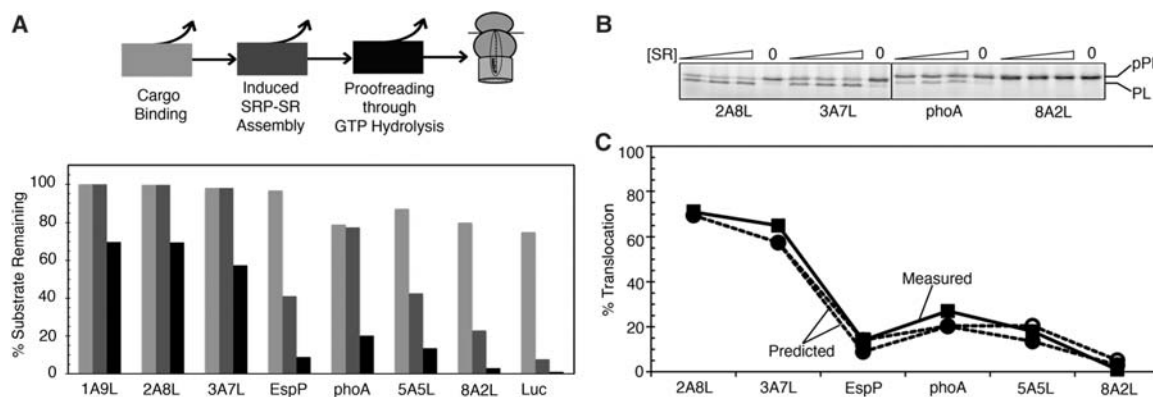


Fig. 3. Correct cargos accelerate GTP-dependent complex formation but delay GTP hydrolysis. (A and B) Rate constants of SRP-SR complex assembly in the nonhydrolyzable GTP analog GMPPNP measured by FRET. (A) RNC_{1A9L} (●) and RNC_{2A8L} (■); (B) RNC_{3A7L} (●), RNC_{phoA} (■), RNC_{5A5L} (◆), and $RNC_{luciferase}$ (▲). Values for k_{cat} were derived from linear fits of data and are summarized in part (C) and table S1. Error bars are SDs from three independent experiments. (C) Summary of GTP-dependent complex assembly rate constants with different

cargos. Error bars are SDs from three independent experiments. (D and E) Effects of cargo on GTP hydrolysis from the SRP-SR complex. (D) Without cargo (●), RNC_{1A9L} (■), and RNC_{5A5L} (◆). (E) RNC_{EspP} (■) and $RNC_{luciferase}$ (◆). Values of k_{cat} were obtained from Michaelis-Menten fits of data and are summarized in (F) and table S1. Error bars are SDs from three independent experiments. (F) Summary of GTPase rate constants in the presence of different cargos. Error bars are SDs from three independent experiments.

Fig. 4. Stepwise rejection of incorrect cargos from the SRP pathway. (A) (Top) Cargos are either retained (straight arrows) or rejected (upward curved arrows) during each checkpoint. (Bottom) Predicted fraction of cargos retained in the SRP pathway during each checkpoint (SOM text). (B) SRP-dependent protein targeting and translocation of the model substrates. pPL and PL



denote the precursor and processed forms of the substrate, respectively. (C) Predicted protein-targeting efficiencies (● and ○) agree well with the experimentally determined values (■), quantified from the data in (B). Trans-

lation elongation rates of 20 (●) and 10 amino acids per second (○) were used for the *E. coli* and eukaryotic ribosomes, respectively, to calculate the targeting efficiencies.

checkpoints to ensure fidelity have been demonstrated in transfer RNA synthetases (27), protein synthesis (28), and DNA and RNA polymerases (29, 30) and possibly represent a general principle for complex cellular pathways that need to recognize degenerate signals or to discriminate between correct and incorrect substrates based on minor differences.

References and Notes

1. P. Walter, A. E. Johnson, *Annu. Rev. Cell Biol.* **10**, 87 (1994).
2. T. A. Rapoport, *Nature* **450**, 663 (2007).
3. M. Halic *et al.*, *Nature* **444**, 507 (2006).
4. M. Halic *et al.*, *Science* **312**, 745 (2006).
5. C. Schaffitzel *et al.*, *Nature* **444**, 503 (2006).
6. M. R. Pool, J. Stumm, T. A. Fulga, I. Sinning, B. Dobberstein, *Science* **297**, 1345 (2002).
7. X. Zhang, S. Kung, S. O. Shan, *J. Mol. Biol.* **381**, 581 (2008).
8. X. Zhang, C. Schaffitzel, N. Ban, S. O. Shan, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1754 (2009).
9. T. Connolly, R. Gilmore, *Cell* **57**, 599 (1989).
10. D. Huber *et al.*, *J. Bacteriol.* **187**, 2983 (2005).
11. J. W. De Gier, Q. A. Valent, G. Von Heijne, J. Lührink, *FEBS Lett.* **408**, 1 (1997).
12. Q. A. Valent *et al.*, *EMBO J.* **14**, 5494 (1995).
13. S. K. Doud, M. M. Chou, D. A. Kendall, *Biochemistry* **32**, 1251 (1993).
14. Materials and methods are available as supporting material on Science Online.
15. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
16. J. H. Peterson, R. L. Szabady, H. D. Bernstein, *J. Biol. Chem.* **281**, 9038 (2006).
17. C. Schaffitzel, N. Ban, *J. Struct. Biol.* **158**, 463 (2007).
18. T. Bornemann, J. Jöckel, M. V. Rodnina, W. Wintermeyer, *Nat. Struct. Mol. Biol.* **15**, 494 (2008).
19. C. G. Jensen, S. Pedersen, *J. Bacteriol.* **176**, 7148 (1994).
20. J. J. Flanagan *et al.*, *J. Biol. Chem.* **278**, 18628 (2003).
21. I. Buskiewicz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7902 (2004).
22. N. Zheng, L. M. Gierasch, *Cell* **86**, 849 (1996).
23. J. D. Miller, H. D. Bernstein, P. Walter, *Nature* **367**, 657 (1994).
24. T. Powers, P. Walter, *EMBO J.* **16**, 4880 (1997).
25. S. O. Shan, S. Chandrasekar, P. Walter, *J. Cell Biol.* **178**, 611 (2007).
26. R. S. Hegde, S. W. Kang, *J. Cell Biol.* **182**, 225 (2008).
27. A. R. Fersht, M. M. Kaethner, *Biochemistry* **15**, 3342 (1976).
28. M. V. Rodnina, W. Wintermeyer, *Annu. Rev. Biochem.* **70**, 415 (2001).
29. T. A. Kunkel, K. Bebenek, *Annu. Rev. Biochem.* **69**, 497 (2000).
30. S. M. Uptain, C. M. Kane, M. J. Chamberlin, *Annu. Rev. Biochem.* **66**, 117 (1997).
31. We thank C. Schaffitzel and N. Ban for help with the purification of RNCs and Trigger factor; J. Lührink for plasmids encoding the *phoA* signal sequence variants; L. Randall for the plasmid encoding luciferase; B. Bukau and E. Deuring for the plasmid encoding Trigger factor; F. Nataro for the plasmid encoding EspP; H. D. Bernstein for the strain HDB51; and R. J. Deshaies, A. Varshavsky, W. Zhong, N. Pierce, and the Shan laboratory for comments on the manuscript. This work was supported by NIH grant GM078024, and career awards from the Burroughs Wellcome Foundation, the Henry and Camille Dreyfus Foundation, the Arnold and Mabel Beckman Foundation, and the Packard Foundation to S.S.

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/757/DC1

Materials and Methods

SOM Text

Table S1

Figs. S1 to S9

References

6 January 2010; accepted 18 March 2010

10.1126/science.1186743

Dynamic Ca^{2+} -Dependent Stimulation of Vesicle Fusion by Membrane-Anchored Synaptotagmin 1

Han-Ki Lee,^{1*} Yoosoo Yang,^{3*} Zengliu Su,^{4*} Changbong Hyeon,⁵ Tae-Sun Lee,¹ Hong-Won Lee,¹ Dae-Hyuk Kwon,³ Yeon-Kyun Shin,^{4,6†} Tae-Young Yoon^{1,2†}

In neurons, synaptotagmin 1 (Syt1) is thought to mediate the fusion of synaptic vesicles with the plasma membrane when presynaptic Ca^{2+} levels rise. However, in vitro reconstitution experiments have failed to recapitulate key characteristics of Ca^{2+} -triggered membrane fusion. Using an in vitro single-vesicle fusion assay, we found that membrane-anchored Syt1 enhanced Ca^{2+} sensitivity and fusion speed. This stimulatory activity of membrane-anchored Syt1 dropped as the Ca^{2+} level rose beyond physiological levels. Thus, Syt1 requires the membrane anchor to stimulate vesicle fusion at physiological Ca^{2+} levels and may function as a dynamic presynaptic Ca^{2+} sensor to control the probability of neurotransmitter release.

Intracellular membrane trafficking in eukaryotic cells involves fusion of membrane-bounded compartments (1) and is mediated by the SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptor) proteins

found on both the vesicle (v-SNARE) and target membranes (t-SNARE) (2–4). In neurons, synaptotagmin 1 (Syt1) is a Ca^{2+} sensor that interacts with SNAREs and membranes to mediate synaptic vesicle fusion, triggering synchronous

neurotransmission (3, 5, 6). Proteoliposome fusion mediated by the t- and v-SNARE proteins is a useful in vitro system (7) for dissecting the molecular functions of presynaptic fusion regulators, including Syt1 (5, 6, 8). The soluble Syt1 variant, which has the two cytoplasmic C2 domains but lacks the transmembrane domain, has been the subject of extensive studies (9–14). This “soluble C2AB” accelerates proteoliposome fusion to reach a time constant of about 10 s when 100 μM or 1 mM Ca^{2+} is added (13, 14). However, when the full-length, membrane-anchored Syt1 is used, addition of the same levels

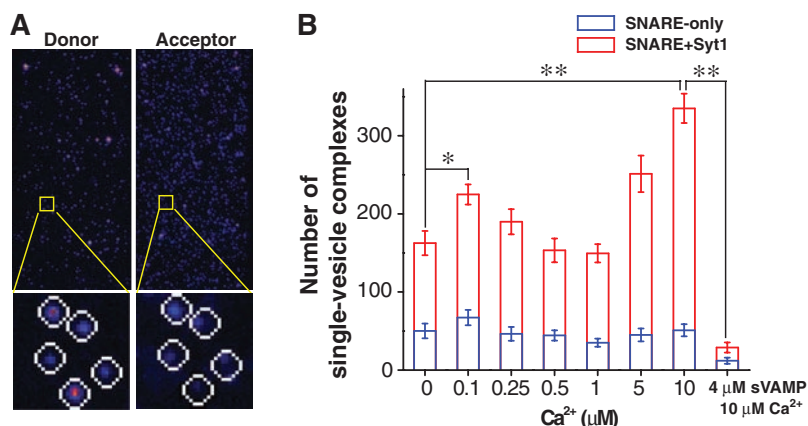
¹Department of Physics, KAIST, Daejeon 305-701, South Korea.

²Institute for the BioCentury, KAIST, Daejeon 305-701, South Korea.

³Department of Genetic Engineering, Sungkyunkwan University, Suwon, Gyeonggi-do 4400-746, South Korea. ⁴Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, 4152 Molecular Biology Building, Ames, IA 50011, USA. ⁵School of Computational Sciences, Korea Institute for Advanced Study, Seoul 130-722, South Korea. ⁶Division of Integrative Biosciences and Biotechnology, POSTECH, Pohang 790-784, South Korea.

*These authors contributed equally to this work. †To whom correspondence should be addressed. E-mail: colishin@iastate.edu (Y.-K.S.); tyoon@kaist.ac.kr (T.-Y.Y.)

Fig. 1. Membrane-anchored Syt1 stimulates single-vesicle docking using 10 μM Ca^{2+} . (A) Exemplary images of single-vesicle FRET imaging. Fluorescence signals from single-vesicle complexes were separated with the threshold at 645 nm and detected as the donor and the acceptor channel signals, respectively (top). Each single-vesicle complex appearing as a Gaussian peak is identified (bottom). (B) Number of single-vesicle complexes formed between t- and v-vesicles under the conditions depicted. The surface-immobilized v-vesicles contain both v-SNARE and Syt1 (red bar) or only v-SNAREs (blue bar). * $P < 0.05$ and ** $P < 0.01$, assessed using the paired t test, and all the errors are SD unless otherwise specified.



This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of May 6, 2010):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/328/5979/760>

Supporting Online Material can be found at:

<http://www.sciencemag.org/cgi/content/full/328/5979/760/DC1>

This article **cites 29 articles**, 6 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/328/5979/760#otherarticles>

This article appears in the following **subject collections**:

Cell Biology

http://www.sciencemag.org/cgi/collection/cell_biol

of Ca^{2+} instead inhibits fusion (12, 15, 16). This membrane anchor is preserved in all isoforms of the Syt family except Syt17 (17), suggesting that it plays an important role in regulating synaptic

vesicle fusion. Understanding this currently hidden role of the membrane anchor requires the study of Ca^{2+} -evoked vesicle fusion stimulated by membrane-anchored Syt1.

Here, we used a single-vesicle fluorescence fusion assay to track the time course of individual vesicle-vesicle fusion events (fig. S1) (18, 19). The assay involves one group of vesicles (called

Fig. 2. Membrane-anchored Syt1 catalyzes full fusion on the hundreds of ms scale in response to $10 \mu\text{M}$ Ca^{2+} . (A to D) Exemplary real-time traces of single-vesicle fusion events. (Top) The changes in the donor (blue) and acceptor (red) fluorescence intensities. (Bottom) The corresponding changes in FRET (black) and stepwise increases in the FRET signals identified by Schwarz information criterion (orange). (E) Cumulative distributions of ΔT for different molecular conditions. (F) Fitting of the ΔT distributions shown in (E) using two exponential functions: $A_1[1 - \exp(-t/\tau_1)] + A_2[1 - \exp(-t/\tau_2)]$. (G) Acceleration of different fusion steps by membrane-anchored Syt1 in response to $10 \mu\text{M}$ Ca^{2+} (SOM text 1). Each parameter was normalized by that of $0 \mu\text{M}$ Ca^{2+} reaction.

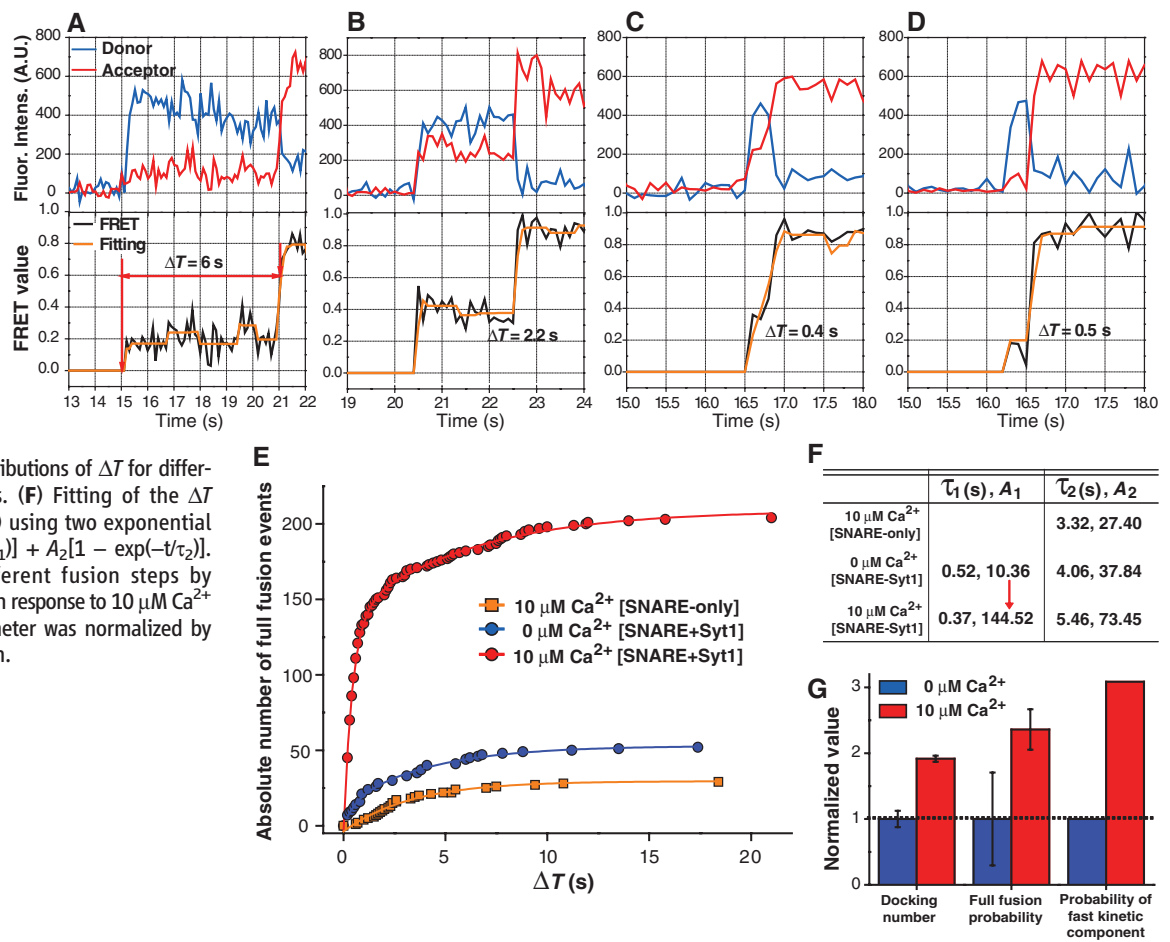
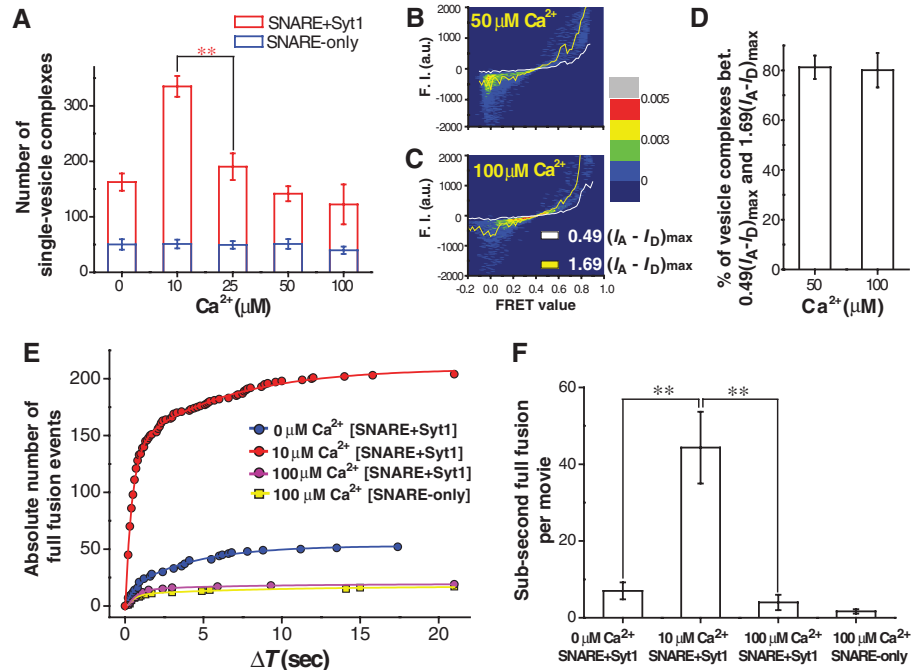


Fig. 3. Membrane-anchored Syt1 is inactivated at sub-mM Ca^{2+} levels. (A) Docking-number analysis for the Ca^{2+} range between 10 and $100 \mu\text{M}$ Ca^{2+} . (B to D) Fluorescence intensity analysis of the single-vesicle complexes. Single-vesicle complexes are placed on the graph plane (FRET efficiency, $I_A - I_D$) to make a density plot (SOM text 2). The percentages of single-vesicle complexes between white and yellow lines, $0.49(I_A - I_D)_{\text{max}}$ and $1.69(I_A - I_D)_{\text{max}}$ (rough estimates for 30% difference in size), are 81.2% ($50 \mu\text{M}$) and 80.5% ($100 \mu\text{M}$ Ca^{2+}) (D). (E) Cumulative distributions of ΔT of single-vesicle fusion events for different molecular conditions (each based on three independent real-time movies). The total docking numbers are 1166 (blue), 2232 (red), 1010 (purple), and 1272 (yellow), respectively. (F) The absolute number of subsecond full-fusion events ($\Delta T < 1$ s) per imaging area, for different protein compositions and Ca^{2+} concentrations.



v-vesicles), containing v-SNARE, full-length Syt1 and lipidic acceptor dyes, immobilized on an imaging surface. A second group of vesicles (t-vesicles), with the precomplex of t-SNAREs and lipidic donor dyes, are mixed with a desired concentration of Ca^{2+} and introduced for reaction via microfluidic buffer exchange. When the t-vesicles react with the surface-immobilized v-vesicles, a single-vesicle complex is formed that can be detected by an increase in the efficiency of the fluorescence resonance energy transfer (FRET) between the donor and acceptor dyes. Docking between t- and v-vesicles generates a FRET efficiency of less than 0.2, whereas the fully fused state generates 0.75 or higher (18).

We first analyzed the ability of membrane-anchored Syt1 to stimulate SNARE-mediated single-vesicle docking using μM Ca^{2+} levels. After a 3-s reaction between the t- and v-vesicles, we counted the number of single-vesicle com-

plexes formed per imaging area (Fig. 1A and figs. S2 to S5). In the absence of Ca^{2+} , Syt1 increased the docking number by a factor of three (Fig. 1B). The stimulation of docking by Syt1 was Ca^{2+} -dependent, and there was another two-fold enhancement between 0 and 10 μM Ca^{2+} (Fig. 1B). We also tested whether this Syt1 effect on docking was strictly dependent on SNARE activity. We disabled the t-SNAREs using a SNARE-motif peptide of v-SNARE (sVAMP), which suppressed the docking number by as much as 90% (Fig. 1B). This observation indicates that Syt1's stimulation essentially involves formation of the ternary SNARE complexes (20–22).

To control the late steps of neurotransmitter release, Syt1 should have the capacity to catalyze the transition to the full-fusion state (3, 5, 6). To explore this possibility, we employed a real-time tracking approach, which followed the entire course of fusion events occurring in individual

single-vesicle complexes (Fig. 2, A to D, fig. S6, and movie S1). The time gap between the docking and full fusion, ΔT , exclusively measured the full-fusion kinetics of an individual fusion event, separated from the docking kinetics (Fig. 2, A to D, and SOM text 1). When the v-vesicle contained only v-SNAREs, the cumulative plot of ΔT shows a single exponential distribution [three independent movies used for each plot (Fig. 2, E and F)]. Inclusion of Syt1 without Ca^{2+} (Fig. 2E) created a small, fast-kinetic component (Fig. 2F, $\tau_1 = 520$ ms). When we added 10 μM Ca^{2+} to this reaction with membrane-anchored Syt1 (Fig. 2E), the absolute population of the fast-kinetic component increased by a factor of 14 (Fig. 2F). This increase of the fast component was not simply due to docking enhancement, but rather was the result of direct acceleration of the full-fusion kinetics (Fig. 2G and SOM text 1). Although it was shown that lipid mixing could occur without proper mixing of contents (23), this fast-fusion kinetics provides convincing evidence that we have observed actual fusion events rather than slow lipid rearrangements such as flip-flop transitions. Thus, in response to 10 μM Ca^{2+} , the membrane-anchored Syt1 directly accelerates the full-fusion reaction, in addition to the docking kinetics.

To determine what happens when the Ca^{2+} concentration increases beyond 10 μM , we repeated the docking-number analysis while increasing Ca^{2+} to 100 μM (Fig. 3, A to D). Unexpectedly, the population of single-vesicle complexes that peaked at 10 μM Ca^{2+} dropped as the Ca^{2+} level rose further, which was visible at 25 μM Ca^{2+} (Fig. 3A). At 100 μM Ca^{2+} , the docking number became comparable to that obtained with no Ca^{2+} (Fig. 3A). In addition, the fluorescence intensity analyses excluded the possibility of multiple-vesicle aggregation at high Ca^{2+} concentrations (Fig. 3, B to D, and SOM text 2). We also performed real-time tracking measurements with higher Ca^{2+} and found that 100 μM Ca^{2+} indeed cut down the catalysis of full-fusion states (Fig. 3E); the absolute population showing a full-fusion event was smaller than that of the 0 μM Ca^{2+} reaction and similar to that observed for the SNARE-only reaction without Syt1 (Fig. 3E). Thus, at 100 μM Ca^{2+} , membrane-anchored Syt1 is largely deactivated and loses its capacity to catalyze the full-fusion state (Fig. 3F).

To study the physical mechanism of Syt1 activity (Fig. 4), we weakened the interaction between Syt1 and the cis-membrane by removing negatively charged phosphatidylserine (PS) lipids from the v-vesicle membrane (9, 12, 24). Without the PS lipids in the cis-membrane, the docking stimulation by Syt1 lost the tendency to decrease up to 100 μM Ca^{2+} (Fig. 4A). On the other hand, when we decreased the phosphatidylinositol 4,5-bisphosphate (PIP₂) lipids in the t-vesicle membrane from 6 (25) to 0.5 mole percent (mol %), the stimulatory effect of Syt1 was abolished in the entire Ca^{2+} range studied (Fig. 4A).

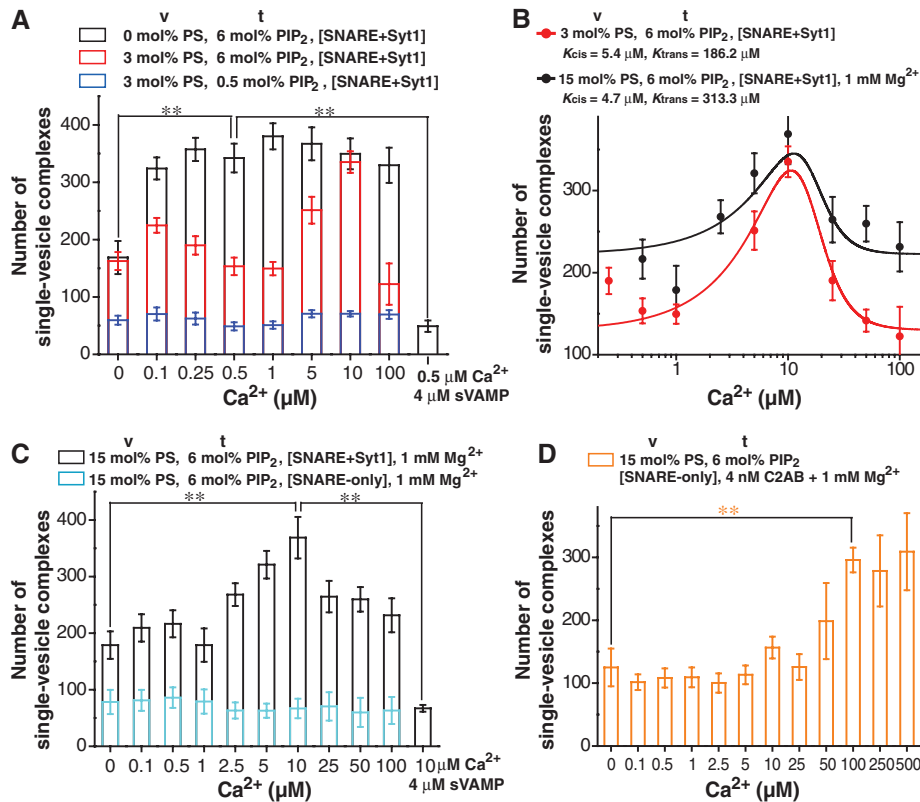


Fig. 4. Dynamic Ca^{2+} -dependent stimulation of vesicle fusion by membrane-anchored Syt1. (A to D) Single-vesicle docking data under the molecular conditions depicted for the cis (v-vesicle) and trans (t-vesicle) membranes. Reactions are carried out with the membrane-anchored Syt1 in the cis-membrane [A to C] or soluble C2AB in the fusion buffer (D) (see SOM text 4). Theoretical fitting of the docking-number data using a modified MWC model (B) (SOM text 3). 15 mol % PS lipids are included for the trans-membrane in every case. (E) Molecular model for the activity of the membrane-anchored Syt1.

Thus, Syt1 needs to interact with both the t-SNARE precomplex (Figs. 1B and 4A, sVAMP treatment) and PIP₂ on the transmembrane for its stimulatory effect, whereas the folding back of Syt1 to the cis-membrane causes its inactivation (Fig. 4E) (12, 24).

The data also suggest that the balance of Syt1's cis- and trans-membrane interactions shifts dynamically as a function of Ca²⁺ concentration. From the single-vesicle data (Fig. 4A), we presume two conformations of Syt1, cis- and trans-conformers, where only the trans-conformer stimulates fusion (SOM text 3 and fig. S7). The Ca²⁺-dependent power shift between these two conformers could be explained by assigning different Ca²⁺ dissociation constants to the cis- (K_{cis}) and trans-conformers (K_{trans}) (Fig. 4B), reminiscent of the classical Monod-Wyman-Changeux (MWC) model for protein allostery. Our model predicts a large anisotropy in the two dissociation constants; $K_{\text{trans}}/K_{\text{cis}} \approx 34$ (Fig. 4B), indicating that at high Ca²⁺ levels where $[\text{Ca}^{2+}]/K_{\text{cis}} \gg 1$, Syt1 predominantly partitions into the cis-conformer and becomes inactivated (Fig. 4E). Because the C2B domain exhibits a much larger Ca²⁺-binding dissociation constant (5, 10, 26), it is tempting to speculate that the trans-conformer, critical for fusion stimulation, is predominantly mediated by the C2B domain.

Finally, we observed that the presence of another divalent ion, Mg²⁺, affected the overall shape of the Ca²⁺-dependent stimulation pattern. Simply increasing the PS lipids in the cis-membrane to the physiological level of 15 mol % (27) pushed the stimulation peak to sub- μM Ca²⁺ levels, followed by a quick deactivation of Syt1 that began at 1 μM Ca²⁺ (fig. S8A). However, in the presence of physiological 1 mM Mg²⁺, the stimulation peak at 10 μM Ca²⁺ was restored (Fig. 4, B and C), probably due to the enhanced screening of electrostatic interactions by Mg²⁺ ions. Furthermore, this nonmonotonic docking pattern correlated well with up- and down-regulation of the ability of Syt1 to accelerate the full-fusion kinetics (fig. S8, B and C). Thus, the essential molecular activities of membrane-anchored Syt1 can be reproduced under physiological charge

conditions in the cis- and trans-membranes. Synaptic vesicles directly isolated from rat brain, which have native Syt1 and lipids, also show a deactivation behavior at 100 μM and 1 mM Ca²⁺ (16). These observations collectively suggest that the dynamic Ca²⁺-dependent activity of Syt1 would occur in physiological contexts.

We have demonstrated that anchoring Syt1 to a lipid membrane fundamentally changes its molecular activity. In the presynaptic active zone, tens of μM or even several μM of presynaptic Ca²⁺ is sufficient to trigger strong neurotransmitter release (28, 29). The membrane-anchored Syt1 reconstituted in our assay has a remarkable Ca²⁺ sensitivity, responding to those Ca²⁺ levels operational at the presynaptic termini. In comparison, at the same effective concentration, we found that the soluble C2AB required one order of magnitude higher Ca²⁺ for docking stimulation and did not show any deactivation pattern up to 500 μM Ca²⁺ (SOM text 4). Thus, in addition to the enhanced Ca²⁺ sensitivity, Syt1's membrane anchor seems to give it the Ca²⁺-dependent, nonmonotonic activity. A simple use of sub-mM Ca²⁺ only inactivates membrane-anchored Syt1, which may explain why previous in vitro studies failed to observe stimulatory effects with full-length Syt1. Our work suggests that Syt1 modulates the probability of neurotransmitter release in response to the presynaptic Ca²⁺ levels and molecular composition of membranes, which may contribute to the dynamic plasticity of neuronal communication.

Reference and Notes

1. J. S. Bonifacio, B. S. Glick, *Cell* **116**, 153 (2004).
2. J. E. Rothman, *Nature* **372**, 55 (1994).
3. T. C. Sudhof, *Annu. Rev. Neurosci.* **27**, 509 (2004).
4. R. Jahn, R. H. Scheller, *Nat. Rev. Mol. Cell Biol.* **7**, 631 (2006).
5. E. R. Chapman, *Annu. Rev. Biochem.* **77**, 615 (2008).
6. J. Rizo, C. Rosenmund, *Nat. Struct. Mol. Biol.* **15**, 665 (2008).
7. T. Weber et al., *Cell* **92**, 759 (1998).
8. A. T. Brunger, K. Weninger, M. Bowen, S. Chu, *Annu. Rev. Biochem.* **78**, 903 (2009).
9. W. C. Tucker, T. Weber, E. R. Chapman, *Science* **304**, 435 (2004).
10. D. Araç et al., *Nat. Struct. Mol. Biol.* **13**, 209 (2006).

11. S. Martens, M. M. Kozlov, H. T. McMahon, *Science* **316**, 1205 (2007).
12. A. Stein, A. Radhakrishnan, D. Riedel, D. Fasshauer, R. Jahn, *Nat. Struct. Mol. Biol.* **14**, 904 (2007).
13. M. Xue, C. Ma, T. K. Craig, C. Rosenmund, J. Rizo, *Nat. Struct. Mol. Biol.* **15**, 1160 (2008).
14. E. Hui, C. P. Johnson, J. Yao, F. M. Dunning, E. R. Chapman, *Cell* **138**, 709 (2009).
15. L. K. Mahal, S. M. Sequeira, J. M. Gureasko, T. H. Söllner, *J. Cell Biol.* **158**, 273 (2002).
16. M. Holt, D. Riedel, A. Stein, C. Schuette, R. Jahn, *Curr. Biol.* **18**, 715 (2008).
17. M. Craxton, *BMC Genomics* **8**, 259 (2007).
18. T. Y. Yoon, B. Okumus, F. Zhang, Y. K. Shin, T. Ha, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 19731 (2006).
19. T. Y. Yoon et al., *Nat. Struct. Mol. Biol.* **15**, 707 (2008).
20. X. Zhang, M. J. Kim-Miller, M. Fukuda, J. A. Kowalchuk, T. F. Martin, *Neuron* **34**, 599 (2002).
21. A. Bhalla, M. C. Chicka, W. C. Tucker, E. R. Chapman, *Nat. Struct. Mol. Biol.* **13**, 323 (2006).
22. H. de Wit et al., *Cell* **138**, 935 (2009).
23. Y. H. Chan, B. van Lengerich, S. G. Boxer, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 979 (2009).
24. J. Bai, W. C. Tucker, E. R. Chapman, *Nat. Struct. Mol. Biol.* **11**, 36 (2004).
25. D. J. James, C. Khodthong, J. A. Kowalchuk, T. F. Martin, *J. Cell Biol.* **182**, 355 (2009).
26. E. Hui, J. Bai, E. R. Chapman, *Biophys. J.* **91**, 1767 (2006).
27. S. Takamori et al., *Cell* **127**, 831 (2006).
28. R. Schneggenburger, E. Neher, *Curr. Opin. Neurobiol.* **15**, 266 (2005).
29. J. Sun et al., *Nature* **450**, 676 (2007).
30. T.-Y.Y. thanks W. Bae and J. Ryu for writing analysis programs, K. Kim for help with the supporting movie, and T. Ha and W. Jones for critical reading of the manuscript. This work was supported by Korea Research Foundation grants (KRF-2008-313-C00365 to T.-Y.Y.; KRF-C00142 and KRF-C00180 to C.H.), and National Research Foundation of Korea grants funded by the Korean government (2009-0087691 to T.-Y.Y.; R01-2008-000-10920-0 to C.H.; 2007-D00243 to D.-H.K.). This work was also supported by a National Institutes of Health grant (R01 GM051290 to Y.-K.S.) and the World Class University program in Korea (R31-2008-000-10105-0 to Y.-K.S. and R33-10163 to T.-Y.Y.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/328/5979/760/DC1
Materials and Methods
SOM Text
Figs. S1 to S9
Table S1
Movie S1
References

1 February 2010; accepted 22 March 2010
10.1126/science.1187722